4-18-2018 3:30 PM

Computational and Analytical Modelling of Droplet-Macroion Interactions

Myong In Oh
The University of Western Ontario

Supervisor
Styliani, Constas
The University of Western Ontario

Graduate Program in Chemistry
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

Recommended Citation

Oh, Myong In, "Computational and Analytical Modelling of Droplet-Macroion Interactions" (2018). *Electronic Thesis and Dissertation Repository*. 5297.
https://ir.lib.uwo.ca/etd/5297

Abstract

Charged droplets involving macromolecules undergo distinct disintegration mechanisms and shape deformations as a consequence of droplet-macroion interactions. Three general classes of droplet-macroion interactions that have been identified in the Consta group are: contiguous extrusion of a linear macroion from a droplet, "pearl-necklace" droplet conformations, and "star"-shaped droplets. This dissertation probes in a systematic manner the onset and various outcomes of macroion-droplet interactions, using atomistic molecular dynamics and realistic examples of solvent and macromolecules.

When the charge-squared-to-volume ratio of a droplet is below but near a threshold value, certain flexible macromolecules, such as poly(ethylene glycol), extrude from a droplet, induced by the charging of the macromolecules. An analytical model is constructed based on the simulation data to suggest that the droplet surface electric field may play a role in the extrusion of the macroion. The effect of different solvents is studied to show that the final charge state of the macroion is determined by complicated macromolecule-ion-solvent interactions.

Beyond this threshold, the charge-induced instability evolves to certain droplet deformations that lead to new stable states. These include "pear-shaped" lobes of solvent at the termini of a linear macroion, such as unstructured proteins, and conical protrusions of dielectric solvent surrounding a macroion regardless of its shape. In the former, such droplet conformation may emerge due to the interplay of a number of factors, subject to the constraint that each sub-droplet should be below a certain charge-squared-to-volume ratio. In the latter, the overall star geometry is determined by the amount of the macroion charge.

As the next level of system complexity, different factors that affect the stability of weak transient protein complexes in droplets are examined. A multiscale approach is devised to model a protein in an evaporating droplet where its acidity constantly changes. A methodology is then developed to compute the dissociation rate and the error in the dissociation constant measured in mass spectrometry experiments. A possible charging mechanism of the macroion due to the "star" structure of solvent is also proposed.

Keywords: Macroion-solvent interactions, Electrostatic interactions, Charging mechanisms, Protein complexes, Charge-induced instabilities, Droplets, Rayleigh limit, Equilibrium constant, Shape fluctuations, Molecular dynamics, Electrospray ionization mass spectrometry, Dynamics, Evaporation

Co-Authorship Statement

Chapter 2 is based on the articles "**Oh, M. I.**, Consta, S. Charging and release mechanisms of flexible macromolecules in droplets. *J. Am. Soc. Mass Spectrom.* **28**, 2262–2279 (2017)" and "Soltani, S., **Oh, M. I.**, Consta, S. Effect of solvent on the charging mechanisms of poly(ethylene glycol) in droplets. *J. Chem. Phys.* **142**, 114307 (2015)", and the submitted manuscript "**Oh, M. I.**, Consta, S. Exploring the extrusion mechanisms of proteins from droplets and of "droplets from macroions". *Anal. Chem.* submitted on February 28, 2018. Manuscript ID: ac-2018-00943b". All the work presented in Chapter 2 was performed by M. I. Oh. The analytical model was developed both by M. I. Oh and S. Consta. S. Consta assisted with the revision of the chapter. Any work performed by S. Soltani (graduate student who initiated her PhD in 2013) is not included in the thesis.

Chapter 3 is based on the manuscript in preparation, entitled "Effect of the dielectric properties and molecular structure of solvents on the formation of star-shaped droplets". All the work presented in Chapter 3 was performed by M. I. Oh. The analytical model was developed both by M. I. Oh and S. Consta. S. Consta assisted with the derivation of the extended Rayleigh model and the revision of the chapter.

Chapter 4 is based on the articles "**Oh, M. I.**, Consta, S. Stability of a transient protein complex in a charged aqueous droplet with variable pH. *J. Phys. Chem. Lett.* **8**, 80–85 (2017)", "**Oh, M. I.**, Consta, S. What factors determine the stability of a weak protein-protein interaction in a charged aqueous droplet? *Phys. Chem. Chem. Phys.* **19**, 31965–31981 (2017)", and "Consta, S., Sharawy, M., **Oh, M. I.**, Malevanets, A. Advances in modeling the stability of noncovalent complexes in charged droplets with applications in electrospray ionization-MS experiments. *Anal. Chem.* **89**, 8192–8202 (2017)". All the work presented in Chapter 4 was performed by M. I. Oh. The analytical model was developed both by M. I. Oh and S. Consta. S. Consta assisted with the revision of the chapter. Any work performed by M. Sharawy and A. Malevanets is not included in the thesis.

Acknowlegements

I would like to take this opportunity to express my deepest gratitude to Prof. Styliani Consta for her supervision and patient guidance on my research.

I thank the current and former group members in the Consta group - Mahmoud Sharawy, Sepideh Soltani, Victor Kwan, Chloe Graham, Ryan Szukalo, and Dr. Maxim Paliy - for their time and contributions to my research.

I thank Prof. Daan Frenkel for granting me a once-in-a-lifetime opportunity to work in the Frenkel group at the University of Cambridge. I also thank Dr. Anatoly Malevanets for familiarizing myself with NAMD and Monte Carlo methods.

I appreciate, from the bottom of my heart, sincere support and earnest prayer from my dearest family - my mother Jung Soo Lim, my sister Young In Oh, and my twin brother Chang In Oh. I would like to give special thanks to my brother since he is the very first mathematician who shepherded me to the beautiful meadow of pure mathematics. Also I thank (and miss) my two puppies, Jadoo (*plum*) and Hodoo (*walnut*), who gladly dedicated their time to relieving my mental stress. Lastly, I thank Jesus Christ for blessing me and my family much more than we deserve.

Contents

List of Figures

List of Tables

List of Appendices

List of Abbreviations

BSA - Buried Surface Area
CEM - Charge Ejection Model
CGenFF - CHARMM General Force Field
CHARMM - Chemistry at Harvard Macromolecular Mechanics
CI-EM - Charge-Induced Extrusion Mechanism
COM - Centre Of Mass
CRM - Charged Residue Model
DMSO - Dimethyl Sulfoxide
DNA - Deoxyribonucleic Acid
ESI - Electrospray Ionization
IEM - Ion Evaporation Mechanism
IM-MS - Ion Mobility-Mass Spectrometry
IMS-MS - Ion Mobility Separation-Mass Spectrometry
LJ - Lennard-Jones
MC - Monte Carlo
MD - Molecular Dynamics
MM - Molecular Mechanics
MS - Mass Spectrometry
NAMD - Nanoscale Molecular Dynamics
NMR - Nuclear Magnetic Resonance
OPLS-AA - Optimized Potential for Liquid Simulations - All Atom
PBC - Periodic Boundary Condition
PDB - Protein Data Bank
PEG - Poly(Ethylene Glycol)
PPI - Protein-Protein Interaction
QM - Quantum Mechanics
RMSD - Root-Mean-Square Displacement
SASA - Solvent Accessible Surface Area
TIP3P - Transferable Intermolecular Potential with Three Points

TRC	-	Transfer Reaction Coordinate
Ub	-	Ubiquitin
UbA	-	Ubiquitin-Associated
UbB	-	Ubiquitin-Binding
VMD	-	Visual Molecular Dynamics

List of Symbols

a - Fitting parameter
a_{l,m_l} - Expansion coefficient of spherical harmonics Y_{l,m_l}
A - Surface area of a droplet
A_{int} - Interfacial area of a protein complex
B - Physical observable
C - Constant
$C(\mathrm{pH}, \mathrm{pK}_a)$ - Population of the protonated state of an amino acid
d - Density
D - Diameter
$\boldsymbol{D}$ - Electric displacement
e - Elementary charge
E - Total energy of a droplet
$\boldsymbol{E}$ - Electric field
E_{diel} - Electrostatic energy of a linear dielectric with free charge
E_{surf} - Surface energy of a droplet
E_{Coul} - Electrostatic energy of a conducting droplet
δE - Energy perturbation
f_{bound} - Binding frequency of specifically interacting particles
F - Electrostatic force
$\boldsymbol{F}_i$ - Net force acting on particle i
$\Delta G^{\ddagger}$ - Activation free energy barrier
h - Planck constant
H - Cone height
I - Ionic strength
k - Rate constant
K_a - Acid dissociation constant
k_B - Boltzmann constant
K_d - Equilibrium constant of dissociation
K_w - Equilibrium constant for water autoprotolysis

L - Contour length of a linear macromolecule

l - Degree of spherical harmonics

m_i - Mass of particle i

m_l - Order of spherical harmonics

M - Molecular mass

$\boldsymbol{n}$ - Outward normal vector to a surface

$\bar{n}_{HB}$ - Time-average number of hydrogen bonds

N - Number of molecules or ions

N_A - Avogadro constant

N_{CF} - Number of Coulomb fissions

p - Pressure

$\boldsymbol{p}$ - Linear momentum

$P_{1/2}$ - Legendre polynomial of order 1/2

$P(t)$ - Distribution of complex fragmentation time

q - Charge

q' - Image charge

Q - Charge

Q_I - Corrected Rayleigh limit for dielectrics

Q_p - Protein charge

Q_R - Rayleigh limit

r - Radius or distance

r' - Distance of an image charge

r_c - Cutoff distance

r_{CM} - Separation distance between two centres of mass

r_s - Switch distance

R_0 - First term ($l = 0$) of the expansion of ρ in terms of spherical harmonics

R - Radius of a droplet

R_g - Radius of gyration

R_{EE} - End-to-end distance

s - Dimensionality of space

S - Total number of time steps

t - Time

T - Temperature

u_{ij} - Pairwise interaction potential

U - Potential energy

V - Volume

W_e - Electrostatic energy between a conductor and a point charge

X - Fissility parameter

Y_{l,m_l} - Spherical harmonics of degree l and order m_l

z_i - Charge of ionic species i

Z_{avg} - Average charge state of a protein

β - Reciprocal of the Boltzmann constant times temperature

χ - Time-average of the number of water molecules in the first few solvation shells per solvent accessible surface area for each monomer

χ_{MeCN} - Mole fraction of acetonitrile

η - Fraction of protein complex population that undergoes dissociation

∇ - Del

ε_0 - Vacuum permittivity

ε_{ext} - Electric permittivity of an external medium

ε_i, ε^I - Relative permittivity of a dielectric

ε_o, ε^E - Relative permittivity of an external medium

ε_{ij} - Well depth of the Lennard-Jones potential

ϵ_s - Depth of a square-well potential

γ - Surface tension

κ - Number of dissociating protein complexes

λ - Length of the extruded portion of a macromolecule from a droplet

ν - Solvation energy per unit length of a macromolecule

Φ - Electrostatic potential

ρ - Shape fluctuation (distance from the centre of a droplet)

σ - Width of a normal distribution

σ_{ij} - Separation distance of particles i and j when the pair Lennard-Jones potential is zero

τ - Complex dissociation rate

τ_r - Length of time elapsed

θ - Polar angle

θ_0 - Complement of the Taylor angle

ϕ - Azimuthal angle

ζ_i - Activity coefficient of ionic species i

$\in$ - Included in a set

Chapter 1

Introduction

1.1 Motivation

The study of the interactions of single monovalent and multivalent ions with water molecules in finite-size systems has been a highly active area of research over several decades in both computations [1–3] and experiments [4–8]. The considerable volume of research in this field is outside the scope of this thesis, so the referenced articles are only indicative and certainly constitute only a small sample. Those studies have focused on clusters comprising several to a few tens of solvent molecules and a single ionic species with charge +1 e to +3 e (where $e = 1.602 \times 10^{-19}$ C is the elementary charge) or pairs of ionic species. The prime interest of ion-solvent interactions stems from the fact that they are fundamental in life and nature; ions do always exist in water, physiological fluids, and other solutions of salts in any phase of matter. In particular, ion-solvent interactions in finite-size systems are studied in order to understand how the structure of a larger system is grown from its smaller analogue and for direct applications in atmospheric aerosols [9, 10].

Clusters with many simple ions and macroions have been, however, less studied at the molecular level, in comparison with bulk solutions [5, 6, 11, 12] involving ions and small clusters with ions. It is only recently that cryogenic spectroscopy and optical spectroscopies coupled to mass spectrometry allow for the detection of solvent-macroion (*e.g.*, nucleic acid and protein) interactions in clusters. The size of these clusters is determined by the size of the macroion, which may be substantial relative to the amount of the solvent (a typical value of a protein radius is 2 nm) and the surrounding solvent layers. We will use the term “droplet” for a macroion-solvent finite-size system in order to distinguish it from the clusters of several tens of solvent molecules. Multiply charged droplets are ubiquitous as found in thunderstorm clouds, but also atmospheric aerosols [10], ionization techniques in mass spectrometry [13–16], electrospinning [17], and surface coating and deposition techniques [18]. Indeed, the practical applications of multiply charged droplets were recognized early by Sir Geoffrey Ingram Taylor in one of his seminal articles published in 1965, which states that "*the practical interest of the [disintegration of water drops in a strong electric field] is that it seems to be an important factor in the production of thunderstorms at any rate in those parts of them where it is too warm for ice crystals to exist* [19]."

This dissertation examines the solvent-macroion interactions in multi-charged droplets composed of several hundreds to thousands of solvent molecules. In the study, atomistic molecular dynamics methods and analytical modelling are used. The study has direct applications in the electrospray ionization mass spectrometry method [13, 14].

1.2 Historical and Theoretical Background on the Study of Charged Droplets

The phenomenology of charged droplets attracted the attention of the scientists since the dawn of electrostatics. The history of charged droplets dates back to at least the late 16th century when William Gilbert [20] noticed the deformation of a liquid water drop into a cone in the presence of a piece of charged amber. Approximately 150 years later, Jean-Antoine Nollet observed aerosolization of water flowing from an electrified vessel with electrical grounding, as he reported that "*a person, electrified by connection to a high-voltage generator, would not bleed normally if he were to cut himself; blood would spray from the wound* [21]."

In the Cavendish laboratory at the University of Cambridge, several distinguished scientists investigated the spatial distribution of charges in a finite volume and the stability of charged droplets. In 1882, Lord Rayleigh [22] presented in his seminal article the criterion for the stability of a spherical charged conductor that undergoes shape fluctuations. Two decades later, his student J. J. Thomson [23], in addition to the discovery of electrons and the setting

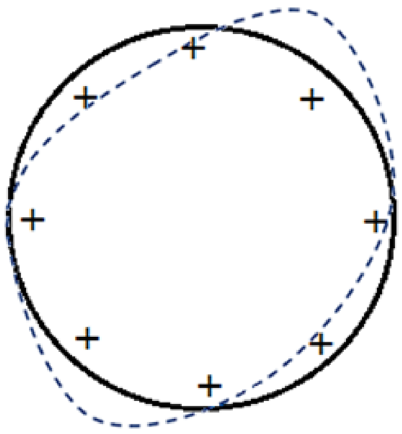


Figure 1.1: Conducting droplet (solid line) undergoing small shape fluctuations (dashed line) relative to a spherical shape. The "+" signs represent charges on the droplet surface.

of the principles of mass spectrometry, posed the so-called Thomson problem to determine the minimum-energy configuration of electrons constrained to be located to the surface of a unit sphere . In 1964, G. I. Taylor pioneered the study of liquid droplet deformations in an electric field [19].

In the following discussion, the theoretical background of Rayleigh's model for a conducting sphere and the formation of a Taylor cone are reviewed. In the Rayleigh model, the total energy of a conducting droplet (E) in a macroscopic (continuum) description is expressed as the sum of the surface energy ($E_{\text{surf}} = \gamma A$) and the electrostatic energy ($E_{\text{Coul}} = \frac{1}{2}Q\Phi$):

$$E = E_{\text{surf}} + E_{\text{Coul}} = \gamma A + \frac{1}{2}Q\Phi \tag{1.1}$$

where Q, γ, and A represent the net charge, the surface tension, and the surface area of the droplet, respectively, and Φ denotes the electrostatic potential on the conductor surface. The model considers that the droplet undergoes small shape fluctuations, where the volume is maintained constant but the surface area may vary. A schematic picture of the fluctuating droplet is shown in Figure 1.1. Since the droplet fluctuates in shape, its surface is given by:

$$\rho(\theta, \phi) = R_0 + \sum_{l>0, m_l} a_{l,m_l} Y_{l,m_l}(\theta, \phi) \tag{1.2}$$

where R_0 is the $l = 0$ term in the expansion, (θ, ϕ) is a set of polar and azimuthal angles, respectively, ρ is the distance from the centre, $Y_{l,m_l}(\theta, \phi)$ denotes the spherical harmonic function of degree l and order m_l, and a_{l,m_l} is the expansion coefficient. The objective of the model is to find the difference in energy (δE) between the perturbed shape and the spherical shape. Following several algebraic steps [24], one arrives at the following expressions for the surface area

$$A = 4\pi R^2 + \sum_{l>0,m_l} \left|a_{l,m_l}\right|^2 + \frac{1}{2} \sum_{l>0,m_l} l(l+1)\left|a_{l,m_l}\right|^2 \tag{1.3}$$

and for the electrostatic potential at the droplet surface

$$\Phi(R) = \frac{1}{4\pi\varepsilon_0}\left[\frac{Q}{R} - \frac{Q}{4\pi R^3} \sum_{l>0,m_l} (l-1)\left|a_{l,m_l}\right|^2\right] \tag{1.4}$$

where ε_0 is the vacuum permittivity. A linear stability analysis of the total energy of the droplet leads to the perturbation energy relative to that of the spherical shape, which is given by

$$\delta E = \frac{1}{2} \sum_{l>0,|m_l|\leq l} (l-1)\left[(l+2)\gamma - \frac{Q^2}{(4\pi)^2\varepsilon_0 R^3}\right]\left|a_{l,m_l}\right|^2 \tag{1.5}$$

The first mode that becomes unstable is found by the first l value that may make one of the coefficients of $\left|a_{l,m_l}\right|^2$ negative. It is found that this is the $l = 2$ mode. The detailed derivation of the Rayleigh expression (Equation 1.5) is shown in Appendix B. Finally, the Rayleigh model gives rise to a dimensionless parameter, which is called the fissility parameter (X) defined as

$$X = \frac{Q^2}{64\pi^2\gamma\varepsilon_0 R^3}. \tag{1.6}$$

In Equation 3.8, all the physical quantities have the same meanings as those in Equations 1.1-1.5. $X = 1$ corresponds to the Rayleigh limit. Hereafter we will denote the charge at the Rayleigh limit by Q_R. Equation 1.5 shows that if X is less than unity, then the system is stable with respect to the $l = 2$ deformations. The droplet state where $X < 1$ is said to be "below the Rayleigh limit"and where $X > 1$ "above the Rayleigh limit". For a charged conducting droplet below the Rayleigh limit, the surface tension, as the restoring force that holds the droplet together, overcomes the electrostatic repulsion, which tends to deform the droplet. We emphasize that the Rayleigh limit defines the onset of instability but not its outcome. Above the critical limit ($X > 1$), the droplet is unstable, therefore it cannot exist in its current spherical or almost spherical form. When a droplet becomes unstable, it will deform or fragment in order to be found in a new stable state.

In a conducting droplet, the instability leads to droplet fission (which is not predicted by the Rayleigh model). The Rayleigh theory provides a remarkably accurate prediction for the charge-induced instability of droplets containing free ions [25–27], even for nanoscopic droplets [28–30].

Also, it is very interesting that the picture of the stability of a charged liquid droplet has been extended to the fission of an atomic nucleus in nuclear physics. The liquid drop model, originally proposed by George Gamow [31] between 1928 and 1936 and further developed by Niels Bohr and John Archibald Wheeler [32] in 1939, treats the nucleus as an incompressible fluid drop of protons and neutrons. In this model, the strong forces that hold nucleons together and the mutual repulsion between protons are analogous to the cohesive force by surface tension and the Coulomb repulsion between charges in a liquid drop, respectively. Therefore, the electrostatic interactions of protons annul to a large extent the effect of the strong nuclear forces that oppose the deformation of the spherical nucleus. Then, in view of the liquid drop model, nuclear fission can be connected to binary Coulomb fission, the process concerned with the splitting of a liquid drop into two smaller droplets, as a consequence of shape deformation caused by an external disturbance [32, 33].

A spherical droplet with a non-fissile ion acquires a “star” shape when its charge-squared-to-volume ratio exceeds a threshold value [34], because charge separation via fission is not allowed in this system. Recently, the Rayleigh model has been extended to predict the onset of charge-induced instability of a dielectric droplet with a central point charge [35]. In this extended Rayleigh model, it is assumed that the electric field from the multiply charged ion is sufficiently weak that the electric field and the resultant polarization have a linear relation [35].

The seminal work of Lord Rayleigh was extended by G. I. Taylor [19] in 1965. G. I. Taylor studied the mechanics of the development of a cone-jet when a small volume of a conducting liquid at the tip of a tube is subjected to a uniform electric field (Figure 1.2). When the electrostatic repulsion begins to counteract the surface tension of the droplet, the hemispherical meniscus is distorted to a conical shape with a semi-vertical angle of 49.3°, which was named

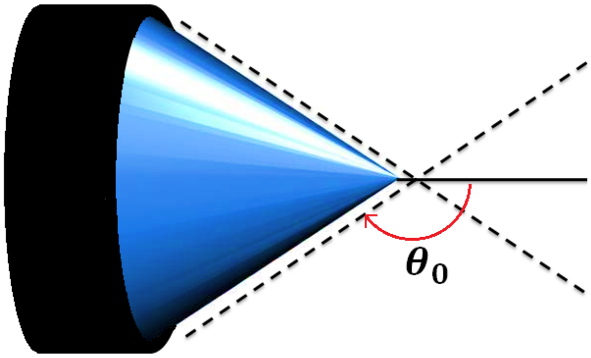


Figure 1.2: Formation of a conical interface of semi-vertical angle $\pi - \theta_0$ when an external electric field is applied to the conducting liquid, coloured in blue, at the tip of a tube, coloured in black.

later a "Taylor cone". Taylor described the static structure of the Taylor cone in which a balance between capillary and electrostatic stresses is fulfilled exactly on the equipotential surface of the cone:

$$\gamma \nabla \cdot \boldsymbol{n} - \frac{1}{2}\varepsilon_0 \boldsymbol{E}^2 = \Delta p \tag{1.7}$$

where γ is the surface tension of the liquid, $\nabla \cdot \boldsymbol{n}$ is the curvature of the cone (∇ is the del operator and $\boldsymbol{n}$ is the normal to the surface directed towards the external medium), ε_0 is the electric permittivity of the external medium (*i.e.*, vacuum), and Δp is the pressure difference. (Vectors will be denoted by letters in boldface.) The solution to Equation 1.7 is the electrostatic potential Φ (which is related to the electric field $\boldsymbol{E}$ by $\boldsymbol{E} = -\nabla\Phi$ and has azimuthal symmetry) in spherical coordinates, expressed in terms of the Legendre polynomial of order $1/2$ ($P_{1/2}$):

$$\Delta\Phi(r, \theta) = C r^{1/2} P_{1/2}(\cos\theta) \tag{1.8}$$

where C is a constant. Substituting Equation 1.8 into Equation 1.7 and demanding the cone surface to be equipotential yields the values of C and θ_0, the complement of the Taylor angle. It turned out that $P_{1/2}(\cos\theta_0) = 0$ and therefore $\theta_0 = 130.7^\circ$ is the only solution for $\theta_0 \in [0, \pi]$. When a certain threshold voltage is reached, the charge accumulated at the apex of the Taylor cone generates an electric field sufficiently strong to overcome the surface tension. Here, the conical vertex inverts and emits a thin jet of liquid removing the excess charge. A Taylor cone is commonly found in electrospraying processes.

Later in 1995, Ramos and Castellanos [36] related the relative permittivity (ε_i) of a dielectric i to the half-angle of the Taylor cone ($\pi - \theta_0$) by the following expression:

$$\frac{\varepsilon_i}{\varepsilon_0} = -\frac{P'_{1/2}(\cos\theta_0)P_{1/2}(-\cos\theta_0)}{P'_{1/2}(-\cos\theta_0)P_{1/2}(\cos\theta_0)} \tag{1.9}$$

where the prime denotes derivative with respect to the argument.

1.3 Macroscopic Models of Droplet Disintegration

One of the most practical applications where charged droplets play a critical role is the electrospray ionization (ESI) method. ESI operates by spraying a solution composed of solvent, macroions, buffer and possibly other additives into a chamber where there is an electric potential difference [37]. The spraying creates a mist of highly charged droplets, which are composed of solvent, macroions and other simpler ions. The droplets undergo a cascade of solvent evaporation and disintegration events [24, 38, 39] that finally lead to desolvated macroions. The dried macroions may be analyzed or selected by mass spectrometry (MS). Since the advance of electrospray ionization mass spectrometry (ESI-MS) in the late 1960s [13, 38, 39], a significant question has been addressed on how macromolecules emerge from carrier droplets to the gas phase at the late stage of an ESI process. The release mechanism of the macroions is directly linked to their final charge states and how these charge states differ from those in bulk solution. Two macroscopic models, the charged-residue model (CRM) [13] and the ion evaporation mechanism (IEM) [38, 39], were proposed to describe different ionization mechanisms that a macromolecule may undergo. These two conventional models form a basis in the development of models of droplet disintegration.

The IEM, also known as the ion desorption model, was first developed by Iribarne and Thomson [38, 39]. In this model, the formation of gaseous ions of low molecular mass was associated with the releases of solvated ions from an evaporating droplet. The IEM states that when droplets shrink to radii of approximately 10 nm, direct emissions of simple ions may take place as an activated process with the rate constant

$$k = \frac{e^{-\beta\Delta G^{\ddagger}}}{\beta h} \tag{1.10}$$

where $\beta = 1/k_BT$ is the reciprocal of the Boltzmann constant (k_B) times temperature (T), h is the Planck constant, and $\Delta G^{\ddagger}$ is the height of the activation free energy barrier. In this mechanism, small solvated ions detach from the parent droplet below the Rayleigh limit due to the electric field emanating from the droplet surface. Therefore, in the regime of droplet radii less than 10 nm, Coulomb fission is superseded by ion evaporation as a pathway to reduce excess charge. The IEM is based on the macroscopic description of the escape event, where the activation free energy is estimated by using the Born model [40] and the transition state theory in reaction kinetics. The transition state of ion evaporation is found where the long-range repulsion between the leaving ion and the remaining charges on the droplet is compensated by the short-range attraction between the leaving ion and the polarized droplet. The IEM does not, however, account for the droplet shape fluctuations at the transition state which are responsible for the fragmentation event [28, 29, 41].

Experimental evidence that validates the IEM was provided by Fernández de la Mora and his co-workers [42–44]. Theoretical treatments of ion evaporation from charged droplets and ion fission from charged clusters [28, 29, 45–47] were also performed by using molecular simulations. The former employed molecular dynamics (MD) simulations on nano-sized charged droplets to understand their structure and fission dynamics. Direct MD calculations are inefficient in capturing the rare event of ion evaporation over the course of droplet evolution. Several MD simulations do not enable one to locate the transition state and the activation barrier of the process and thus the rate constant of the fragmentation process of the droplet. The limitations of MD become more pronounced for the system at low temperature and/or well below the Rayleigh limit due to the higher rarity of the process. One of the methods to overcome this sampling issue is to perform an infinitely large number of different initial conditions of the system, which is not possible in practice. Therefore, Consta *et al.* [28, 29, 45–47] devised a novel collective reaction coordinate, called the transfer reaction coordinate, that couples the degrees of freedom of ions and solvent. They performed Monte Carlo simulations in which one can identify the shape fluctuations that are critical for the escape of the ions and then estimate the rate of the process. This methodology samples the configurations of the system at the transition state, and therefore, it resolves the problem of observing only a small number of ion release events in a direct MD run.

The charge-residue model (CRM) was proposed by Dole *et al.* [13] to describe the formation of gaseous ions of an analyte of high molecular mass in an ESI process. In this model, very small droplets (of radius $\approx$ 1 nm) that form when solvent evaporation nears completion harbour one macromolecule as well as charge carriers on the droplet surface. Complete desolvation will then lead to the charged residue formation of the macromolecule in the gas phase via the landing of the ions on the macromolecule. The significance of the CRM is often found in the prediction of the final charge states of electrosprayed proteins and their complexes. The key equation to this model is the correlation, empirically obtained by Tolić *et al.* [48] and theoretically derived by Fernández de la Mora [49], between the average charge states Z_{avg} and the molecular masses M of globular proteins:

$$Z_{avg} = 8\pi\left(\frac{\varepsilon_0\gamma R^3 M}{e^2}\right)^{1/2} = \left(\frac{48\pi\varepsilon_0\gamma M}{e^2 d N_A}\right)^{1/2} = 0.078M^{1/2} \tag{1.11}$$

where R is the radius of the protein (or the final "droplet"), γ and d are the surface tension and the density of water, respectively, and N_A is the Avogadro constant. In his derivation, the following critical assumptions were made: (1) The ultimate aqueous droplet (generated at the latest stage of the droplet evolution) is slightly larger in size than the globular protein molecule that it contains; (2) The final solvent evaporation transfers the droplet residual charges completely to the protein; (3) The protein would be neutral if water disappears so that the charges on droplet surface become the charge of the protein observed in ESI-MS; (4) The protein and water have the same density; (5) Evaporating charged droplets always stay close to the Rayleigh limit. The experimental findings were within 90-110% of the Rayleigh prediction in Ref. [50] for the common solvents of water, acetonitrile, and methanol. The discrepancy between the theoretical prediction and the experimental results has been ascribed to the different charge distribution within an electrosprayed droplet, relative to that of a droplet in a vacuum, due to the electric field created by the neighbouring droplets in an electrospray plume [51].

Consequently, this empirical relation (*i.e.*, Equation 1.11) enables quantitative predictions of the charge states of gaseous globular proteins and their complexes [52, 53] in ESI-MS. In this model, however, chemical reactions that take place in the presence of other solution additives are not considered. For instance, it has been known that the protonation state of an electrosprayed protein may change due to supercharging [54–57] or decharging by the addition of bases [58], crown ethers [59], and cations with large sizes and low solvation energies [60].

In 2008, Hogan *et al.* [61, 62] proposed a modified model of the CRM where evaporation of small ions precedes collapse of remaining ions onto a protein. The kinetics of small ion emission at the critical electric field determines the charge state of an evaporating droplet. As solvent dries out, remaining charge carriers are transferred to the macromolecule that is assumed to dwell in the interior of the droplet. Native proteins electrosprayed from aqueous solutions containing different electrolytes support this combined charged residue-field emission model. The authors reported lower charge states of native proteins in ESI mass spectra when salt additives, such as ammonium acetate and triethylammonium bicarbonate, were added in millimolar concentrations in the solution to be electrosprayed. In particular, lower energy will be required for surface active ions to undergo fission from the parent droplet by the critical electric field. The authors remark that their model seems applicable to most proteins in ESI-MS, with the exception of "supercharged" proteins observed in the positive ion mode with highly acidic solutions.

There are a plethora of applications in physical and biological sciences where charged droplet-macromolecule interactions play a pivotal role. In ESI-MS, the understanding of the interactions is required for the correct interpretation of the spectra. Also, ESI-MS is commonly used to measure equilibrium constants of noncovalent complexes [63, 64] and analyze the architecture of macromolecular assemblies [65]. The insight on the interactions may also find their applications in other ionization methods such as thermospray [16], sonic spray [15], and matrix-assisted ionization [66]. Other significant applications include the fabrication of fibrillosomes [67], the manipulation of polymer particle morphology [68], the acceleration of protein-sugar binding [69], and the study of polymer crystallization and dynamics under spatial confinement [70, 71]. Despite of their significance, however, the behaviour of macroions and their complexes in a droplet environment is relatively less studied and understood at the molecular level in comparison to those in bulk solution.

1.4 Molecular Modelling of Charged Droplets

At the molecular level, charged droplets may be distinguished based on the presence of different types of charge carriers: (a) simple ions only, (b) macroions only, and (c) simple ions and macroions. Recent computational research [34, 35, 72–79] has evidenced that how different charge carriers interact with a droplet depends largely upon the nature of the charges it holds. In other words, it has been revealed that the different character of charge-induced instabilities is associated with different approaches and aims for those types of charged systems. If a droplet of moderate to large dielectric constant is charged with multiple small ions, the droplet approximates a conductor. This is supported by experiments [80] and numerical simulations [30, 46, 72]. Here, the type of the simple charge carriers varies from protons to polyatomic ions. All these studies confirm the predictive power of the Rayleigh limit for a conducting droplet. In a charged droplet that contains divisible charge carriers below the Rayleigh limit, the fission, if occurs, is asymmetric and occurs as an activated process. The IEM was developed in that context. The closest theory to the IEM in the computational studies is the treatment of the fragmentation process as an activated process using a collective order parameter that couples the degrees of freedom of ions and solvent molecules [28, 29, 45, 47]. In a charged droplet that contains a single macromolecule and simple ions, as found in the most practical cases where macromolecular globules are electrosprayed, the CRM is commonly accepted as a pathway of macromolecular charging, that is, as a direct result of macromolecule-ion interactions upon solvent evaporation. Fenn and his co-workers [81] considered the IEM as the mechanism for macroion release without a theoretical basis analogous to that of IEM for single ions, due to the complexity of macromolecular structures and charging mechanisms. Later, an analytical model has been developed by Consta *et al.* [72, 75, 76], which predicts the extrusion mechanism and the droplet conditions under which a linear macromolecule may be released from a droplet. Furthermore, their model identifies the activation energy barrier in the extrusion mechanism.

When the charge is localized on a single macroion or a complex of macromolecules as covalently bound charge or noncovalently transferable charges between the macroion and the solvent, there are no models to capture the variety of the charge-induced instability as well as the associated droplet structure and dynamics. Molecular simulations is the only viable path to identify the outcomes of the charge-induced instabilities.

Using molecular simulations, Consta and her co-workers [34, 73–76] have identified three general scenarios of macroion-droplet interactions (Figure 1.3): (1) the contiguous extrusion of the macroion from the droplet [73, 74], (2) the spiky protrusion (or "star" formation) of solvent around the macroion [34], and (3) the pearl-necklace model [34]. The first scenario was found by molecular simulations in the release of sodiated, lithiated, and uniformly charged poly(ethylene glycol) (PEG) from an aqueous droplet. The second scenario was observed in simulations of a polyhistidine and nucleic acids in nano-sized aqueous droplets where water forms highly ordered structure around the macroion. The "spiky" droplet appears after the droplet passes an instability point that exhibits enhanced shape fluctuations. The third scenario was revealed for the first time in simulations of PEG with bound charges in methanol. The charge of each sub-droplet ("pearl") is below the Rayleigh limit.

1.5 Outline of the Thesis

This dissertation investigates macroion-droplet interactions by using atomistic molecular dynamics (MD) methods with the following objectives: (1) to identify new motifs of charge-induced droplet disintegration and deformation in the presence of a macromolecule, (2) to discuss their general features and establish analytical models for the mechanisms, and (3) to resolve existing controversy in the field of ESI-MS and its relevant research areas. (Molecular dynamics methods are explained in Appendix C.) The study focuses on nanoscopic droplets of aqueous and organic solvents that solvate charge carriers extending from simple monatomic ions (such as metal cations and halide anions) to polyelectrolytes (such as proteins and their complexes).

In Chapter 2, the mechanisms of charge-induced disintegration of droplets are examined

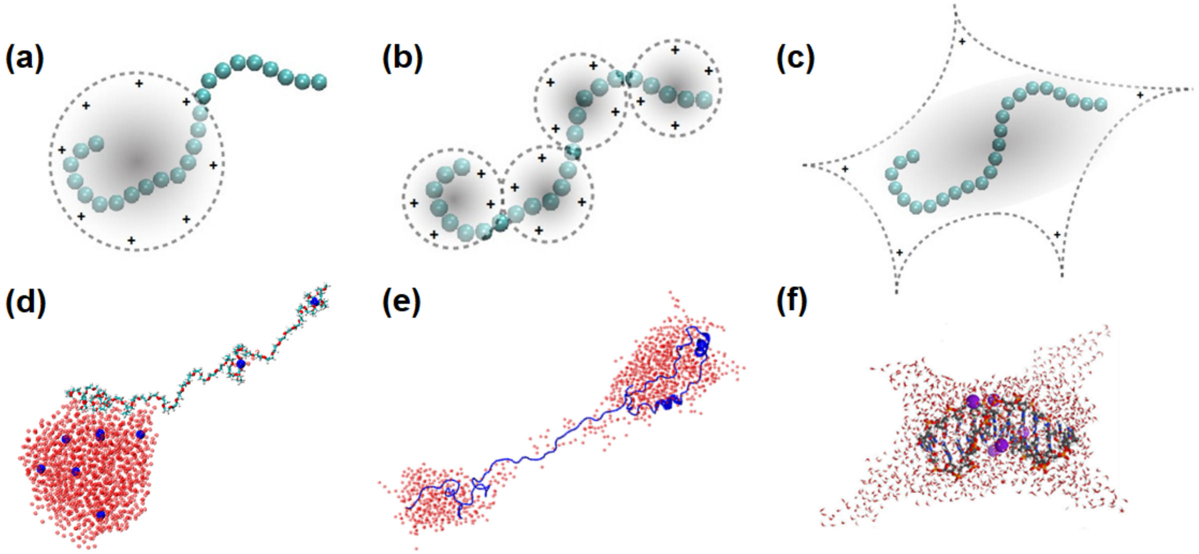


Figure 1.3: Classes of stable droplet-chain conformations. The drawn macromolecule is only a schematic representation of a macroion in a droplet. The solvent boundary is shown by the dashed line. The polarization charge on the droplet surface is shown by positive signs when the macroion is positively charged. (a) Gradual expulsion of a linear macroion from a droplet. This process occurs below the Rayleigh limit. (b) "Pearl-necklace" droplet conformation. The charge within each "pearl" is below the Rayleigh limit. (c) Formation of conical protrusions of the solvent surrounding a macroion. (d) Atomistically modelled extrusion of poly(ethylene glycol) from an aqueous droplet [73] that corresponds to (a). The blue spheres represent Na^+ ions, and the red dots represent H_2O molecules. (e) A droplet comprised approximately 1200 water molecules and $2LJL^{15+}$ [34, 82]. Each spherical lobe is found below the Rayleigh limit as in the schematic (b). (f) A droplet comprised water (red dots), a negatively charged 20mer dsDNA (ds stands for double stranded) and Na^+ ions (purple spheres) [83]. It corresponds to schematic (c).

when they contain flexible linear macromolecules. In the first part of the chapter, by using methylated poly(ethylene glycol) (PEG) as a typical example of flexible synthetic polymer, molecular-level descriptions are provided as to how the macromolecule is ionized and released from a charged droplet of aqueous and organic solvents. Here, a new mechanism of chain extrusion is identified, and an electrostatic model is formulated based on the data obtained from the MD simulations. The significance of this study is found in the fact that there has been long-lasting debates on the manner that the macromolecule emerges from an electrospray droplet and the effect of solvent on the final charge state of the macroion. This part of the chapter aims to resolve these issues by delineating, at the molecular level, how unique solvent-ion-macromolecule interactions affect the conformation and the final charge state of a linear flexible macromolecule. In the second part of the chapter, membrane proteins with a positively charged tail, acid-induced denatured proteins, and intrinsically disordered proteins are used to examine whether they follow the same extrusion mechanism as sodiated PEG from an aqueous droplet. By analyzing protein-water interaction, it appears to be unlikely that common unstructured proteins extrude from a droplet even when the spherical droplet may be found above the Rayleigh limit. Here, a new general motif of macroion-droplet interactions is identified from the consistent findings of the MD simulations. This study contributes in our understanding of linear macroion-droplet interactions.

In Chapter 3, the geometry of “star” morphologies of highly charged dielectric droplets is characterized. The droplets contain a single multiply charged spherical macroion in one of the two different solvents: water and dimethyl sulfoxide (DMSO). We explored how the drastic difference in the solvent properties affects the features of the stellate structures such as the number of rays and the cone angles. By varying the magnitude and the sign of the charge of the macroion, the quadratic relation between the number of rays and the amount of droplet charge is derived and compared with the simulation data. The similarity of the conical shapes with Taylor cones is also examined. The interest in the “star”-shaped droplets arises from (1) their important role in the charge state of electrosprayed macroions, and (2) their potential applications in materials science and catalysis.

As the next level of complexity, the stability of a noncovalent protein complex in an evaporating droplet with variable pH is studied in Chapter 4 by using weak transient dimeric protein complexes. The development of a multi-scale approach is necessary for the atomistic modelling of the system that accounts the protonation state of a protein. A critical issue arises from the fact that the pH of the droplet constantly changes due to solvent evaporation. As simulating proton transfer reactions with direct MD is impractical, it is crucial to devise a special, thorough methodology to resolve this problem. From this study, general features that are applicable to other complexes are extracted, complementing both experimental and computational studies of protein-protein interactions in droplets. This chapter elucidates for the first time the dissociation mechanism of the protein complex, which is significant in experiments (*e.g.*, prevention of noncovalent complexes) and simulations (*e.g.*, order parameters for free energy calculations). This chapter investigates the effects of various factors and the unique role of protein-droplet interactions in terms of charge-induced instabilities, in order to provide molecular-level understanding of stabilization of protein-protein interfaces in the droplet regime. Finally, a mathematical expression is derived from reliable quantities that can be obtained from experiments and simulations, so as to find the error in the measurement of protein binding affinity due to such undesirable complex dissociation in an ESI-MS experiment.

This dissertation is concluded in Chapter 5 with a summary of the core questions posed in this research and the key results.

The majority of the problems posed in the thesis have direct applications to ESI-MS, but it should be noted that this study is not restricted only to this research field but it extends to a vast array of subjects as stated in Section 1.3 that include designing of nanostructures, manipulation of particle morphologies, and controlling noncovalent complex stability. Lastly, every computational methodology and theoretical framework found in the thesis was devised by a collaboration of Dr. Consta and myself. All the Monte Carlo simulation codes (found in Appendix D) and data analysis programs were developed and optimized by myself.

References

[1] Gorlova, O. *et al.* Characterization of the primary hydration shell of the hydroxide ion with H_2 tagging vibrational spectroscopy of the $OH^-\cdot(H_2O)_{n=2,3}$ and $OD^-\cdot(D_2O)_{n=2,3}$ clusters. *J. Chem. Phys.* **145**, 134304 (2016).

[2] Fournier, J. A. *et al.* Snapshots of proton accommodation at a microscopic water surface: Understanding the vibrational spectral signatures of the charge defect in cryogenically cooled $H^+(H_2O)_{n=2-28}$ clusters. *J. Phys. Chem. A* **119**, 9425–9440 (2015).

[3] Perera, L. & Berkowitz, M. L. Stabilization energies of Cl^-, Br^-, and I^- ions in water clusters. *J. Chem. Phys.* **99**, 4222–4224 (1993).

[4] Miller, D. J. & Lisy, J. M. Hydrated alkali-metal cations: Infrared spectroscopy and ab initio calculations of $M^+(H_2O)_{x=2-5}$ Ar cluster ions for M = Li, Na, K, and Cs. *J. Am. Chem. Soc.* **130**, 15381–15392 (2008).

[5] Wang, X.-B. Cluster model studies of anion and molecular specificities via electrospray ionization photoelectron spectroscopy. *J. Phys. Chem. A* **121**, 1389–1401 (2017).

[6] Ross, M. W., Berkdemir, C. & Castleman Jr, A. Strong-field ionization and Coulomb explosion of chlorine weakly bound to small water clusters. *J. Phys. Chem. A* **116**, 8530–8538 (2012).

[7] Chakrabarty, S. & Williams, E. R. The effect of halide and iodate anions on the hydrogen-bonding network of water in aqueous nanodrops. *Phys. Chem. Chem. Phys.* **18**, 25483–25490 (2016).

[8] Wolke, C. T. *et al.* Isotopomer-selective spectra of a single intact H_2O molecule in the $Cs^+(D_2O)_5H_2O$ isotopologue: Going beyond pattern recognition to harvest the structural information encoded in vibrational spectra. *J. Chem. Phys.* **144**, 074305 (2016).

[9] Després, V. R. *et al.* Primary biological aerosol particles in the atmosphere: a review. *Tellus B Chem. Phys. Meteorol.* **64**, 15598 (2012).

[10] Davis, E. & Bridges, M. The Rayleigh limit of charge revisited: light scattering from exploding droplets. *J. Aerosol Sci.* **25**, 1179–1199 (1994).

[11] Lynden-Bell, R. M., Rasaiah, J. C. & Noworyta, J. P. Using simulation to study solvation in water. *Pure Appl. Chem.* **73**, 1721–1731 (2001).

[12] Bagchi, B. & Jana, B. Solvation dynamics in dipolar liquids. *Chem. Soc. Rev.* **39**, 1936–1954 (2010).

[13] Dole, M. *et al.* Molecular beams of macroions. *J. Chem. Phys.* **49**, 2240–2249 (1968).

[14] Wong, S. F., Meng, C. K. & Fenn, J. B. Multiple charging in electrospray ionization of poly(ethylene glycols). *J. Phys. Chem.* **92**, 546–550 (1988).

[15] Hirabayashi, A., Sakairi, M. & Koizumi, H. Sonic spray ionization method for atmospheric pressure ionization mass spectrometry. *Anal. Chem.* **66**, 4557–4559 (1994).

[16] Blakley, C. R. & Vestal, M. L. Thermospray interface for liquid chromatography/mass spectrometry. *Anal. Chem.* **55**, 750–754 (1983).

[17] Bhardwaj, N. & Kundu, S. C. Electrospinning: A fascinating fiber fabrication technique. *Biotechnol. Adv.* **28**, 325–347 (2010).

[18] Salata, O. V. Tools of nanotechnology: Electrospray. *Curr. Nanosci.* **1**, 25–33 (2005).

[19] Taylor, G. Disintegration of water drops in an electric field. *Proc. Royal Soc. A* **280**, 383–397 (1964).

[20] Gilbert, W. *De Magnete*. Dover Classics of Science and Mathematics (Dover Publications, 1958).

[21] Grimm, R. L. *Fundamental Studies of the Mechanisms and Applications of Field-Induced Droplet Ionization Mass Spectrometry and Electrospray Mass Spectrometry*. Ph.D. thesis, California Institute of Technology (2006).

[22] Rayleigh, L. XX. on the equilibrium of liquid conducting masses charged with electricity. *Philos. Mag.* **14**, 184–186 (1882).

[23] Thomson, J. XXIV. on the structure of the atom: an investigation of the stability and periods of oscillation of a number of corpuscles arranged at equal intervals around the circumference of a circle; with application of the results to the theory of atomic structure. *Philos. Mag. Ser. 6* **7**, 237–265 (1904).

[24] Consta, S. & Malevanets, A. Disintegration mechanisms of charged nanodroplets: novel systems for applying methods of activated processes. *Mol. Simul.* **41**, 73–85 (2015).

[25] Duft, D., Lebius, H., Huber, B. A., Guet, C. & Leisner, T. Shape oscillations and stability of charged microdroplets. *Phys. Rev. Lett.* **89**, 084503 (2002).

[26] Duft, D., Achtzehn, T., Muller, R., Huber, B. A. & Leisner, T. Coulomb fission: Rayleigh jets from levitated microdroplets. *Nature* **421**, 128–128 (2003).

[27] Achtzehn, T., Müller, R., Duft, D. & Leisner, T. The Coulomb instability of charged microdroplets: dynamics and scaling. *Eur. Phys. J. D* **34**, 311–313 (2005).

[28] Consta, S. Fragmentation reactions of charged aqueous clusters. *J. Mol. Struct. THEOCHEM* **591**, 131–140 (2002).

[29] Consta, S., Mainer, K. R. & Novak, W. Fragmentation mechanisms of aqueous clusters charged with ions. *J. Chem. Phys.* **119**, 10125–10132 (2003).

[30] Sharawy, M. & Consta, S. Effect of counterions on the charging mechanisms of a macromolecule in aqueous nanodrops. *J. Chem. Phys.* **141**, 104321 (2014).

[31] Stuewer, R. H. The origin of the liquid-drop model and the interpretation of nuclear fission. *Perspect. Sci.* **2**, 76–129 (1994).

[32] Bohr, N. & Wheeler, J. A. The mechanism of nuclear fission. *Phys. Rev.* **56**, 426–450 (1939).

[33] Meitner, L. & Frisch, O. R. Disintegration of uranium by neutrons: a new type of nuclear reaction. *Nature* **143**, 239–240 (1939).

[34] Consta, S. Manifestation of Rayleigh instability in droplets containing multiply charged macroions. *J. Phys. Chem. B* **114**, 5263–5268 (2010).

[35] Oh, M. I., Malevanets, A., Paliy, M., Frenkel, D. & Consta, S. When droplets become stars: charged dielectric droplets beyond the Rayleigh limit. *Soft Matter* **13**, 8781–8795 (2017).

[36] Ramos, A. & Castellanos, A. Conical points in liquid-liquid interfaces subjected to electric fields. *Phys. Lett. A* **184**, 268–272 (1994).

[37] Kebarle, P. & Verkerk, U. H. Electrospray: from ions in solution to ions in the gas phase, what we know now. *Mass Spectrom. Rev.* **28**, 898–917 (2009).

[38] Iribarne, J. & Thomson, B. Evaporation of small ions from charged droplets. *J. Chem. Phys.* **64**, 2287–2294 (1976).

[39] Thomson, B. & Iribarne, J. Field-induced ion evaporation from liquid surfaces at atmospheric-pressure. *J. Chem. Phys.* **71**, 4451–4463 (1979).

[40] Atkins, P. W. & MacDermott, A. J. The Born equation and ionic solvation. *J. Chem. Educ.* **59**, 359 (1982).

[41] Labowsky, M., Fenn, J. & Fernández de la Mora, J. A continuum model for ion evaporation from a drop: effect of curvature and charge on ion solvation energy. *Anal. Chim. Acta* **406**, 105–118 (2000).

[42] Gamero-Castaño, M. & Fernández de la Mora, J. Kinetics of small ion evaporation from the charge and mass distribution of multiply charged clusters in electrosprays. *J. Mass Spectrom.* **35**, 790–803 (2000).

[43] Gamero-Castaño, M. & Fernández de la Mora, J. Direct measurement of ion evaporation kinetics from electrified liquid surfaces. *J. Chem. Phys.* **113**, 815–832 (2000).

[44] Loscertales, I. G. & Fernández de la Mora, J. Experiments on the kinetics of field evaporation of small ions from droplets. *J. Chem. Phys.* **103**, 5041–5060 (1995).

[45] Consta, S., Oh, M. I. & Soltani, S. Advances in the theoretical and molecular simulation studies of the ion chemistry in droplets. *Int. J. Mass Spectrom.* **377**, 557–567 (2015).

[46] Ichiki, K. & Consta, S. Disintegration mechanisms of charged aqueous nanodroplets studied by simulations and analytical models. *J. Phys. Chem. B* **110**, 19168–19175 (2006).

[47] Consta, S. Detecting reaction pathways and computing reaction rates in condensed phase. *Theor. Chem. Acc.* **116**, 373–382 (2006).

[48] Tolić, L. P. *et al.* Electrospray ionization fourier transform ion cyclotron resonance mass spectrometric characterization of high molecular mass starburst (TM) dendrimers. *Int. J. Mass Spectrom. Ion Process.* **165-166**, 405–418 (1997).

[49] Fernández de la Mora, J. Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Anal. Chim. Acta* **406**, 93–104 (2000).

[50] Smith, J. N., Flagan, R. C. & Beauchamp, J. L. Droplet evaporation and discharge dynamics in electrospray ionization. *J. Phys. Chem. A* **106**, 9957–9967 (2002).

[51] Shrimpton, J. S. Dielectric charged drop break-up at sub-Rayleigh limit conditions. *IEEE Trans. Dielectr. Electr. Insul.* **12**, 573–578 (2005).

[52] Heck, A. J. R. & van den Heuvel, R. H. H. Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom. Rev.* **23**, 368–389 (2004).

[53] Mehmood, S., Allison, T. M. & Robinson, C. V. Mass spectrometry of protein complexes: From origins to applications. *Annu. Rev. Phys. Chem.* **66**, 453–474 (2015).

[54] Lomeli, S. H., Yin, S., Loo, R. R. O. & Loo, J. A. Increasing charge while preserving noncovalent protein complexes for ESI-MS. *J. Am. Soc. Mass. Spectrom.* **20**, 593–596 (2009).

[55] Lomeli, S. H., Peng, I. X., Yin, S., Ogorzalek Loo, R. R. & Loo, J. A. New reagents for increasing ESI multiple charging of proteins and protein complexes. *J. Am. Soc. Mass Spectrom.* **21**, 127–131 (2010).

[56] Sterling, H. J. & Williams, E. R. Origin of supercharging in electrospray ionization of noncovalent complexes from aqueous solution. *J. Am. Soc. Mass Spectrom.* **20**, 1933–1943 (2009).

[57] Ogorzalek Loo, R. R., Lakshmanan, R. & Loo, J. A. What protein charging (and supercharging) reveal about the mechanism of electrospray ionization. *J. Am. Soc. Mass Spectrom.* **25**, 1675–1693 (2014).

[58] Catalina, M. I., van den Heuvel, R. H. H., van Duijn, E. & Heck, A. J. R. Decharging of globular proteins and protein complexes in electrospray. *Chem. Eur. J.* **11**, 960–968 (2005).

[59] Pagel, K., Hyung, S.-J., Ruotolo, B. T. & Robinson, C. V. Alternate dissociation pathways identified in charge-reduced protein complex ions. *Anal. Chem.* **82**, 5363–5372 (2010).

[60] Susa, A. C., Mortensen, D. N. & Williams, E. R. Effects of cations on protein and peptide charging in electrospray ionization from aqueous solutions. *J. Am. Soc. Mass Spectrom.* **25**, 918–927 (2014).

[61] Hogan, C. J., Carroll, J. A., Rohrs, H. W., Biswas, P. & Gross, M. L. Charge carrier field emission determines the number of charges on native state proteins in electrospray ionization. *J. Am. Chem. Soc.* **130**, 6926–6927 (2008).

[62] Hogan, C. J., Carroll, J. A., Rohrs, H. W., Biswas, P. & Gross, M. L. Combined charged residue-field emission model of macromolecular electrospray ionization. *Anal. Chem.* **81**, 369–377 (2009).

[63] Boeri Erba, E., Barylyuk, K., Yang, Y. & Zenobi, R. Quantifying protein-protein interactions within noncovalent complexes using electrospray ionization mass spectrometry. *Anal. Chem.* **83**, 9251–9259 (2011).

[64] Barylyuk, K., Gülbakan, B., Xie, X. & Zenobi, R. DNA oligonucleotides: A model system with tunable binding strength to study monomer-dimer equilibria with electrospray ionization-mass spectrometry. *Anal. Chem.* **85**, 11902–11912 (2013).

[65] Rajabi, K., Ashcroft, A. E. & Radford, S. E. Mass spectrometric methods to analyze the structural organization of macromolecular complexes. *Methods* **89**, 13–21 (2015).

[66] Devereaux, Z. J. *et al.* Matrix-assisted ionization on a portable mass spectrometer: Analysis directly from biological and synthetic materials. *Anal. Chem.* **88**, 10831–10836 (2016).

[67] Song, Y. *et al.* Fabrication of fibrillosomes from droplets stabilized by protein nanofibrils at all-aqueous interfaces. *Nat. Commun.* **7**, 12934 (2016).

[68] Almería, B., Deng, W., Fahmy, T. M. & Gomez, A. Controlling the morphology of electrospray-generated PLGA microparticles for drug delivery. *J. Colloid Interface Sci.* **343**, 125–133 (2010).

[69] Lee, J. K., Banerjee, S., Nam, H. G. & Zare, R. N. Acceleration of reaction in charged microdroplets. *Q. Rev. Biophys.* **48**, 437–444 (2015).

[70] Michell, R. M., Blaszczyk-Lezak, I., Mijangos, C. & Müller, A. J. Confinement effects on polymer crystallization: From droplets to alumina nanopores. *Polymer* **54**, 4059–4077 (2013).

[71] Yoshida, K., Horii, K., Saito, A., Takashima, A. & Nishio, I. Confinement effects on polymer dynamics: Thermo-responsive behaviours of hydroxypropyl cellulose polymers in phospholipid-coated droplets (water-in-oil emulsion). *Polymers* **9**, 680 (2017).

[72] Consta, S., Oh, M. I. & Malevanets, A. New mechanisms of macroion-induced disintegration of charged droplets. *Chem. Phys. Lett.* **663**, 1–12 (2016).

[73] Consta, S. & Chung, J. K. Charge-induced conformational changes of PEG-$(Na^+)_n$ in a vacuum and aqueous nanodroplets. *J. Phys. Chem. B* **115**, 10447–10455 (2011).

[74] Chung, J. K. & Consta, S. Release mechanisms of poly(ethylene glycol) macroions from aqueous charged nanodroplets. *J. Phys. Chem. B* **116**, 5777–5785 (2012).

[75] Consta, S. & Malevanets, A. Manifestations of charge induced instability in droplets effected by charged macromolecules. *Phys. Rev. Lett.* **109**, 148301 (2012).

[76] Consta, S. & Malevanets, A. Classification of the ejection mechanisms of charged macromolecules from liquid droplets. *J. Chem. Phys.* **138**, 044314 (2013).

[77] Oh, M. I. & Consta, S. What factors determine the stability of a weak protein-protein interaction in a charged aqueous droplet? *Phys. Chem. Chem. Phys.* **19**, 31965–31981 (2017).

[78] Oh, M. I. & Consta, S. Stability of a transient protein complex in a charged aqueous droplet with variable pH. *J. Phys. Chem. Lett.* **8**, 80–85 (2017).

[79] Sharawy, M. & Consta, S. Characterization of ”star” droplet morphologies induced by charged macromolecules. *J. Phys. Chem. A* **120**, 8871–8880 (2016).

[80] Li, K.-Y., Tu, H. & Ray, A. K. Charge limits on droplets during evaporation. *Langmuir* **21**, 3786–3794 (2005).

[81] Nguyen, S. & Fenn, J. B. Gas-phase ions of solute species from charged droplets of solutions. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 1111–1117 (2007).

[82] Singarapu, K. K. *et al.* Structural characterization of Hsp12, the heat shock protein from saccharomyces cerevisiae, in aqueous solution where it is intrinsically disordered and in detergent micelles where it is locally α-helical. *J. Biol. Chem.* **286**, 43447–43453 (2011).

[83] Sharawy, M. & Consta, S. How do non-covalent complexes dissociate in droplets? A case study of the desolvation of dsDNA from a charged aqueous nanodrop. *Phys. Chem. Chem. Phys.* **17**, 25550–25562 (2015).

Chapter 2

Droplet-Linear Flexible Macroion Interactions

2.1 Motivation and Objectives

The key question in electrospray ionization mass spectrometry (ESI-MS) experiments is the origin and the magnitude of the charge state of a macroion that is detected in the gaseous state by MS [1]. It has been demonstrated in experiments that the droplet environment plays a decisive role in the charging of a macroion [2–4]. Therefore, the understanding of the macroion-droplet interactions is significant for explaining the charging mechanism and the final charge state of a macroion.

The charging mechanism of globular macroions is reviewed in this paragraph to contrast it to that of a linear macroion. One of the broadly accepted mechanisms for the charging of globular proteins in droplets is the charged-residue model (CRM) proposed by de la Mora and his co-workers [4]. The CRM assumes that when a droplet comprises the compact macroion and a thin hydration layer surrounding the macroion in the latest stage of droplet evolution, the droplet holds as many charges as predicted by the corresponding Rayleigh limit (Equation 1.6) This model states that further solvent evaporation will collapse the small ions on the surface of

---

Reproduced by permission of the American Chemical Society from Oh, M. I., Consta, S. "Exploring the extrusion mechanisms of proteins from droplets and of "droplets from macroions"" *Anal. Chem.* submitted on February 28, 2018. Manuscript ID: ac-2018-00943b.

Reproduced by permission of Springer Nature from Oh, M. I., Consta, S. "Charging and release mechanisms of flexible macromolecules in droplets" *J. Am. Soc. Mass Spectrom.* **28**, 2262–2279 (2017).

Reproduced by permission of the American Institute of Physics Publishing from Soltani, S., Oh, M. I., Consta, S. "Effect of solvent on the charging mechanisms of poly(ethylene glycol) in droplets" *J. Chem. Phys.* **142**, 114307 (2015).

the globular protein, and therefore, the final charge of the protein will be equal to the amount of charge predicted by the Rayleigh limit for a liquid droplet of the size of the protein. This model assumes explicitly that (1) the mass density of the protein is the same as that of water, and implicitly that (2) the last layer of water surrounding the macroion is uniformly spread over the surface of the protein. The second assumption may not be always true due to the inhomogeneous composition of amino acids on a protein surface. The CRM is close agreement with the experimentally measured charge states of globular proteins and their complexes [4–7]. In reality, however, there is no direct evidence of the CRM besides it. Consta and her co-workers [8] have proposed a different mechanism from that proposed by de la Mora [4] to explain the final charge state of a globular protein. In this mechanism, the final charge state of the protein is obtained by the emission of protons from the droplet of a higher charge state, which was acquired at an earlier stage of the droplet lifetime because of the acidity conditions. This mechanism also predicts the final charge state of a globular protein that is close to the Rayleigh limit of the protein-sized droplet. This mechanism has been supported by analytical theory [8] and numerical modelling [8], but still needs to be probed in experiments.

The CRM does not suffice to capture all the manners in which a broad spectrum of macromolecules are ionized and released from droplets. In particular, the relation between the final charge state and the geometry of linear polymers has been addressed by Fenn [9, 10] and de la Mora [4]. Typical examples of such macromolecules are flexible linear macromolecules, including poly(ethylene glycol) (PEG), and less flexible macromolecules such as denatured proteins and intrinsically disordered proteins (IDPs). Experiments based only on the interpretation of ESI mass spectra and mobility measurements, were able to associate extended conformations of macromolecules with final charge states much higher than the limiting charges predicted by the Rayleigh limit of their compact analogues. Fenn and de la Mora have also suggested [4, 9, 10] as follows: “*As the effective surface tension of solvent becomes too small to maintain the spherical shape of the droplet, the macroion becomes extended by Coulombic repulsion and, as a result, the droplet becomes elongated with conical ends that accumulate the charge density. Solvent evaporation then continues to leave a linear multiply charged macromolecule* [10].” Although their suggestions as to how droplet morphologies influence macroion charging have been insightful, direct evidence completely lacked at that time. Recently, molecular simulations have shown a more complicated picture of the droplet morphology than that of the two conical ends [11, 12] (see also Section 4).

The objective of this study is to connect complicated ion-solvent-macroion interactions in a droplet to the release mechanism and the final charge state of the macromolecule, by performing molecular simulations with atomistic details. In the first section of this chapter, the effect of solvent on charging and release mechanisms of PEG is examined. To supplement this study, Appendix A analyzes the solvation of PEG in a neutral droplet of a pristine solvent. This section is also relevant to one of the most controversial topics in the field of ESI-MS, namely, the role of surface tension in the final charge state of an electrosprayed macromolecule [13, 14].

We selected to study two solvents, water and acetonitrile. The interactions of PEG with these two solvents place the PEG in different regions in a droplet, that is, on the droplet surface in water *vs.* in the interior of the droplet in acetonitrile. The solvation behaviour of PEG (thus the dimension of its conformation) in a droplet is then manipulated by altering the composition of the binary mixture of water and acetonitrile. PEG has a persistence length of 3.7−3.8 Å [15, 16]. Its persistence length is small relative to the contour length of the specific PEG64 (64 denotes the number of monomers) that we examine, which is approximately 275 Å. Hence, the PEG molecule represents a flexible linear macromolecule. The sodiation of PEG is studied because (a) the interactions of PEG with water and Na^+ ions can be modelled reliably with the existing force fields [17, 18], and (b) PEG-ion interactions have been extensively studied in experiments by the methods of mass spectrometry [4, 19–24].

The next question to be addressed is: “Do protein ions that lack an ordered three-dimensional structure in water droplets follow the same extrusion mechanism as found in the sodiated PEG in aqueous droplets?” This will be answered in the second section of this chapter. It has been reported in the literature that unfolded myoglobin in charge state +17 [25, 26] and intrinsically disordered proteins (IDPs) [27] extrude from water droplets in a similar manner to that of PEG [17, 18, 28]. This study focuses on three examples of proteins, each selected from the broad classes of membrane proteins, IDPs, and denatured proteins. Most membrane proteins feature in their composition a high percentage of non-polar amino acids and low charge. Therefore, a membrane protein resides on the surface of a charged droplet. The composition of IDPs is characterized by a low content of bulky hydrophobic amino acids and a high proportion of polar and charged amino acids [29–32]. Acid-induced denaturation often causes unfolded proteins with high charge, as all basic residues and many acidic residues are protonated under acidic conditions. Therefore, IDPs and denatured proteins are likely solvated in the interior of an aqueous droplet. The manner in which the protein-droplet systems evolve when they are close to the Rayleigh limit will be examined and compared with the findings from the study of the PEG-droplet systems.

The denatured and disordered proteins used in this study contain 100−150 residues. The persistence length is on the order of a nanometer, which corresponds to the size of a single peptide unit [33]. Therefore, these proteins may be classified as relatively flexible macromolecules, but still not as flexible as a PEG molecule due to the presence of large residue groups and the peptide backbone.

2.2 Charging and Extrusion Mechanisms of PEG from a Droplet

2.2.1 Modelling and Computational Methods

Atomistic molecular dynamics (MD) simulations were performed of nano-sized droplets composed of solvent, a methylated poly(ethylene glycol) molecule with 64 monomers (PEG64), and Na^+ ions. The solvent was selected to be H_2O, acetonitrile (MeCN), and a binary mixture of H_2O-MeCN with various compositions. These solvents are commonly used in ESI-MS experiments. PEG was chosen as a realistic model for a flexible linear macromolecule, and Na^+ ions were used as the charge carriers. The PEG and MeCN molecules were modelled using the OPLS-AA (Optimized Potential for Liquid Simulations - All Atom) [34, 35] force field where all hydrogen atoms are explicitly represented. The OPLS-AA force field includes optimized potential terms for bond-stretching and angle-bending vibrations, which are described by an ideal harmonic oscillator, as well as torsional strains, Coulombic electrostatic interactions, and Lennard-Jones interactions. The parameters for the atomic sites of the PEG molecule were taken from those for ethers. The water molecules were represented by the TIP3P model [36] (three-site transferable intermolecular potential). All the MD simulations were carried out using GROMACS [37] version 4.5.5. Visual Molecular Dynamics (VMD) [38] was used for the purpose of all visualizations.

For every charged droplet system, constant-temperature MD simulations were performed using the leap-frog algorithm and velocity rescaling to thermostat the systems and maintain the desired temperature. The time step of integration was set to 0.5 fs in all the runs, and electrostatic and van der Waals interactions were not truncated. Each droplet system was placed *in vacuo* without the periodic boundary condition (PBC) for non-equilibrium runs where solvent and ion evaporation were allowed. The droplets decreased in size by solvent evaporation and release of solvated ions during the simulations. The evaporated molecules and ions were removed when they were found at a distance from the connected body of the system larger than three times the diameter of the droplet.

To initiate the simulations of charged aqueous droplets, a stretched PEG64 molecule was solvated in a cubic solvent box constructed by 7000 water molecules. Several water molecules were replaced with Na^+ ions in order to introduce excess positive charge into the system. Three cases differing in the number of ions (16, 32, and 128 Na^+) were studied. The temperature was set to $T = 300$ K. Initially, energy minimization was performed in the cubic box and then the box was removed for the actual production runs. In order to test the diffusion rate of PEG *vs.* the sodiation rate of PEG, a simulation of PEG30 in a droplet with 25 Na^+ and 10000 water molecules was also performed.

In the simulations of charged MeCN droplets, four droplets, each composed of 800 MeCN molecules, one PEG64 molecule, and Na^+ ions ($N_{Na^+} = 2, 4, 8$, and 12) were prepared and simulated. The initial conformation of the PEG64 molecule in the droplets was compact, which was obtained from the simulation of a bare PEG64 molecule *in vacuo*. For the purpose of comparison, an electrically neutral droplet of MeCN (*i.e.*, $N_{Na^+} = 0$) was also simulated. In order to investigate the effect of the temperature on the charging mechanism of the macromolecule in the MeCN droplets, the temperature was varied. The systems evolved under four different temperatures, namely, 250 K, 280 K, 300 K, and 320 K, which are all lower than the experimental boiling point of bulk acetonitrile (≈ 355 K) [39]. By visual inspection, it was observed that only the pure MeCN droplet charged with 12 e (where $e = 1.602 \times 10^{-19}$ C is the elementary charge) was above the Rayleigh limit at $T = 280$ K, as it undergoes immediate fissions at the very early stage of the realization. Additionally, in order to study the effect of solvation on the conformational size of the macromolecule, a PEG64 molecule with the maximum number of Na^+ ions that it can have, which is five (PEG64-5Na^+) [40], was solvated inside a MeCN droplet and simulated at $T = 280$ K.

Lastly, for charged H_2O-MeCN droplets, five different solvent compositions with mole fractions of MeCN (χ_{MeCN}), ranging from $\chi_{\mathrm{MeCN}} = 0$ (corresponding to pure H_2O) to $\chi_{\mathrm{MeCN}} = 1.0$ (corresponding to pure MeCN), were examined. The total number of the solvent molecules and the amount of Na^+ ions were maintained the same at the initial setup for every system ($N_{\mathrm{solvent}} = 800$, $N_{\mathrm{Na^+}} = 8$). The temperature was maintained approximately at 280 K.

Detailed information about the different parameters used in building systems and running simulations is summarized in Table 2.1.

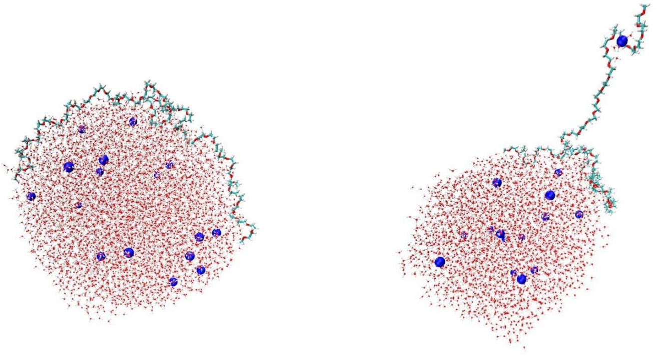
2.2.2 Macromolecules on the Surface of a Water Droplet

Relation of Charging to Release of PEG It was found that, regardless of the contour length of the macromolecule, the initial configuration, and the nature of the charge carriers [18], PEG settles on the surface of an aqueous droplet (see Appendix A for the detailed study on PEG-solvent interactions in a droplet). When confined in a spherical cavity, a flexible macromolecule tends to be centred due to entropic effects [41]. A droplet is a confining environment for a flexible macromolecule, but its solvation in a droplet does not always lead to the central location because enthalpic effects also play a significant role. In this case, the strong hydrogen-bonding network formed by water molecules determines the interfacial confinement of the macromolecule [42–46]. Therefore, the sodiation of PEG takes place when it captures Na^+ ions at the liquid/vapour interface of the droplet.

In the following discussion, the charge states of the droplets in Figure 2.1 are compared with the values predicted by the Rayleigh criterion. The surface tension used is taken from Ref. [47] and Ref. [48] where the surface tension of bulk TIP3P water at $T = 300$ K has been estimated. This value of the surface tension in Ref. [47] was calculated to be 52.3 mN/m with a standard deviation 1.5 and in Ref. [48] 49.5 mN/m. It is known that the presence of sodium halide salts increase the surface tension of water [49], whereas the surface tension of water drops in the presence of PEG as a surfactant [44]. Consequently, the modified surface tension does not change the critical charge estimated by the Rayleigh limit. In the range of maximum

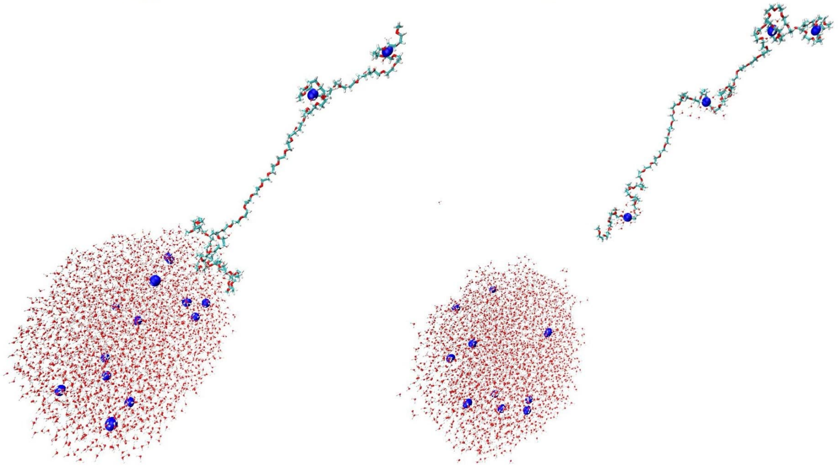
Table 2.1: N_{H_2O}, N_{MeCN}, and N_{Na^+} indicate the number of H_2O and MeCN molecules, and Na^+ ions, respectively. χ_{MeCN} denotes the mole fraction of MeCN in H_2O – MeCN mixture. T is the temperature in Kelvin, and ν_{PEG} is the length of the PEG used. r_s and r_c are the switch and cutoff distances for the switch function. t is the length of the production run.

H_2O **Droplets**								
System	N_{H_2O}	N_{Na^+}	T (K)	ν_{PEG}	PBC box (nm^3)	r_s (nm)	r_c (nm)	t (ns)
1	0	0	300	64	No	-	-	100.0
2	250	0	300	64	Yes (15^3)	4.0	5.0	100.0
3	800	0	300	64	Yes (17^3)	5.0	6.0	100.0
4	1500	0	300	64	Yes (15^3)	5.5	6.5	100.0
5	4000	0	300	64	Yes (21^3)	7.5	8.5	40.0
6	7000	0	300	64	Yes (28^3)	9.0	10.0	11.5
7	4000	0	300	30	Yes (20^3)	7.5	8.5	4.0
8	10000	0	300	30	No	-	-	3.6
9	4000	0	300	40	Yes (20^3)	7.5	8.5	4.0
10	4000	0	300	50	Yes (20^3)	7.5	8.5	4.0
11	4000	0	300	64	Yes (20^3)	7.5	8.5	4.0
12	7000	16	300	64	No	-	-	84.8
13	7000	32	300	64	No	-	-	20.3
14	7000	128	300	64	No	-	-	67.5
MeCN **Droplets**								
System	N_{MeCN}	N_{Na^+}	T (K)	ν_{PEG}	PBC box (nm^3)	r_s (nm)	r_c (nm)	t (ns)
15	0	0	250	64	No	-	-	90.0
16	250	0	250	64	Yes (14^3)	4.5	5.5	105.0
17	800	0	250	64	Yes (17^3)	6.5	7.5	24.0
18	1500	0	250	64	Yes (20^3)	8.0	9.0	8.6
19	800	5 (bound)	280	64	No	-	-	9.0
20	800	2	280	64	No	-	-	10.0
21	800	4	280	64	No	-	-	16.0
22	800	8	280	64	No	-	-	12.0
23	800	12	280	64	No	-	-	12.0
24	800	8	250	64	No	-	-	43.5
25	800	8	300	64	No	-	-	5.0
26	800	8	320	64	No	-	-	3.0
H_2O – MeCN **Droplets**								
System	χ_{MeCN}	N_{Na^+}	T (K)	ν_{PEG}	PBC box (nm^3)	r_s (nm)	r_c (nm)	t (ns)
27	0	8	280	64	No	-	-	30.4
28	0.2	8	280	64	No	-	-	24.9
29	0.5	8	280	64	No	-	-	23.5
30	0.8	8	280	64	No	-	-	10.0
31	1.0	8	280	64	No	-	-	10.0



(a) $N_{H_2O} = 6541$
$N_{Na^+} = 16$

(b) $N_{H_2O} = 3591$
$N_{Na^+} = 15$



(c) $N_{H_2O} = 3476$
$N_{Na^+} = 14$

(d) $N_{H_2O} = 3215$
$N_{Na^+} = 12$

Figure 2.1: Typical snapshots of the release of PEG64 from an aqueous droplet. The colour of the atomic sites is coded as follows: blue spheres represent Na^+ ions, the red sites represent oxygen, the white sites hydrogen and the turquoise carbon. The Na^+ ions have been enlarged relative to the other atomic sites for visualization purpose. For the four configurations, the time t from the initiation of the simulation run, the radius of the droplet (r), the number of H_2O molecules (N_{H_2O}), and the number of Na^+ ions (N_{Na}) remaining in the droplet are written. The duration of the process shown is approximately 65 ns at T = 300 K.

value of surface tension (53.8 mN/m) to minimum (49.5 mN/m), the Rayleigh prediction gives a difference of at most one elementary charge.

Typical snapshots of the charging of PEG64 during the course of desolvation is shown in Figure 2.1. This simulation started from a droplet of 7000 H_2O molecules with 16 Na^+ ions and a PEG64 molecule. These parameters correspond to a charge-squared-to-volume ratio of the droplet considerably below the Rayleigh limit. In Figure 2.1 (a), the droplet is found at 29−31% below the Rayleigh limit. The droplet shrinks in size mainly via solvent evaporation and via the escape of one solvated single Na^+. In Figure 2.1 (b), the droplet has lost almost half of its solvent relative to the initial configuration, and it is approximately 13 % below the Rayleigh limit when the first sodiation of PEG64 occurs. The first sodiated segment of PEG64 is released when the droplet is 10−13% below the Rayleigh limit. In Figure 2.1 (c) and (d), the subsequent sodiations follow and occur 15−18 % and 17−21 % below the Rayleigh limit, respectively. The radius of the droplet remains approximately constant throughout the sodiation and the release of PEG. As shown in Figure 2.1 (b-d), only $\approx$8 % of the water molecules evaporate during the charging and release of the sodiated PEG. Figure 2.2 shows the comparison between the charge state of the droplet when the sodiated PEG is released and the Rayleigh limit corresponding to that particular size of the droplet for the three initial conditions of the simulations. It is clearly seen that the final release of PEG from aqueous droplets occurs below the Rayleigh limit in agreement with experiments [4] and theory [50].

For the simulations with 32 and 128 Na^+ ions, the systems undergo rapid Coulomb fissions at the initial stage until they release all the excess charge. In all the runs, the sodiation of the macromolecule starts close to the Rayleigh limit. In the droplet that carries initially 32 Na^+ and has charge close to the Rayleigh limit, the sodiation of PEG64 occurs early in the droplet's lifetime (see the first set of points in Figure 2.2). For the droplet that is found initially substantially below the Rayleigh limit, the droplet reduces its size first to the square of charge-to-volume ratio close to the Rayleigh limit so as that PEG64 becomes charged (see the intermediate set of points in Figure 2.2). Finally, when the droplet is found substantially above the Rayleigh limit, vigorous Coulomb explosions create a smaller droplet with approximately 2000 water molecules where the macromolecule is found (see the last set of points in Figure 2.2). A common observation made from all these runs is that the sodiation of PEG64 and the release of its

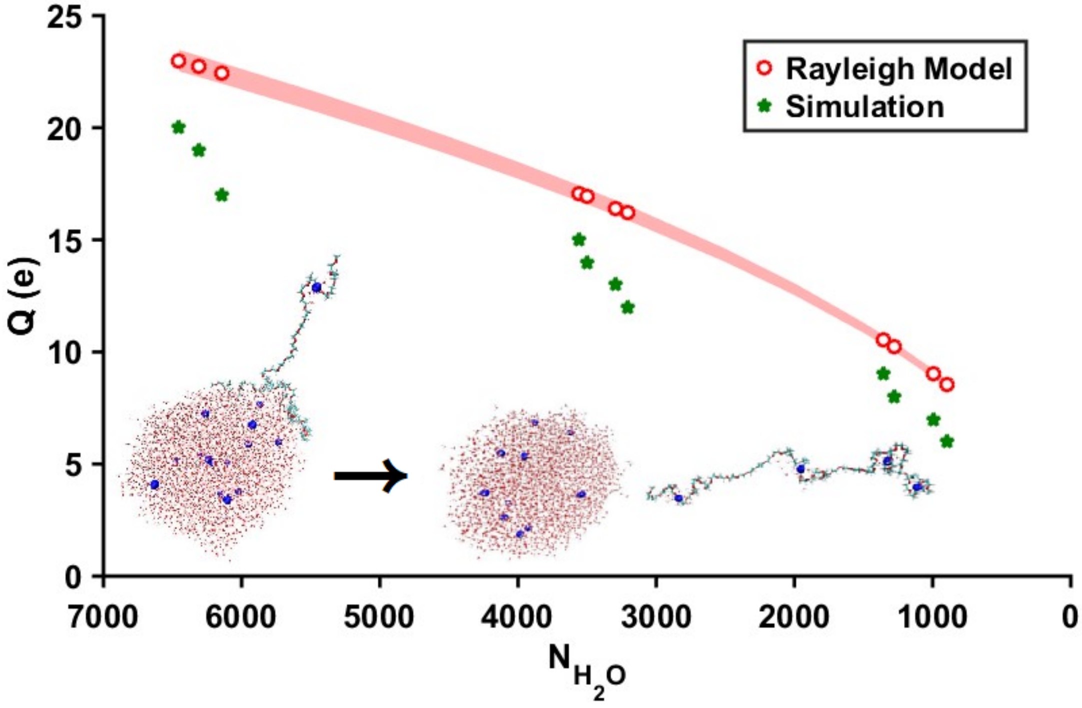


Figure 2.2: Droplet net charge in unit of *e* (where *e* is the elementary charge) *vs*. the number of H_2O molecules remaining in the main body of the droplet as it shrinks by release of solvated Na^+ ions and solvent evaporation. During the droplet lifetime, the PEG64 becomes sodiated and it is released. The snapshots show the release of a sodiated PEG64 out of the droplet. The green stars correspond to the three sets of runs performed at $T = 300$ K starting from different droplet sizes. The net charge of the droplet does not contain the charge that has been captured by the extended segment of PEG64 out of the droplet. The red circles show the theoretical estimates by the Rayleigh limit, and the hazy red region the deviation of the values due to surface tension estimates taken.

first sodiated segment occur at approximately 10−15 % below the Rayleigh limit. Although sodiation may occur fairly below the Rayleigh limit, these Na^+ ions are released back into the aqueous solution of the droplet (see the following discussion). Once the emission of the first sodiated segment of PEG occurs, it triggers the subsequent sodiation of other segments in the regime that gradually deviates from the Rayleigh prediction.

Direct visualization of all the simulations reveals two mechanisms via which the aqueous droplet reduces its excess charge when it is close (from below) to the Rayleigh limit. (i) A Na^+ ion is released by first attaching to PEG. The attachment of the Na^+ ion onto PEG can occur directly by the transfer of the cation from H_2O into PEG where it is wrapped by the ethereal oxygen sites at the surface of the droplet first, and then it is released along with the PEG portion (*i.e.*, the escape of the macromolecule is coupled to its charging mechanism). When a Na^+ ion is located in proximity to PEG, it may be released as a small droplet without initially being wrapped by the macromolecule. The solvated Na^+ ion keeps in contact with PEG as it detaches from the droplet, and it drags a short segment of the PEG with it. Then, the solvated ion rolls on the extruded part of the PEG until the ion is eventually wrapped by PEG when the water shell around the Na^+ ion completely dries out. (ii) The cation escapes the parent droplet simply by a release from the surface of the droplet that is not occupied by PEG. For large droplets, it was found that Na^+ (or $Na(H_2O)_n^+$) emission from the surface occurs rarely compared to Na^+ transfer to PEG within the droplet. The escape of an ion from a conductor such as an aqueous droplet with Na^+ ions is expected to be an activated process because of the induced charge on the conductor. However, in the release of the sodiated segments of PEG, this activation barrier appears to be lower than the release of a water-solvated ion. This is evidenced by the fact that for a certain charge-squared-to-volume ratio in the droplet, the release of the water-solvated ion is rarely observed in comparison to the PEG-coordinated ion. The ion in PEG is surrounded by a hydrophobic sheath that facilitates its release. In principle, one can estimate from the simulations the activation energy by finding the rate of the escape of the water-solvated and PEG-coordinated ions.

The charging of PEG close to the Rayleigh limit can be attributed to the following two possible effects: (i) Close to the Rayleigh limit, a critical density of charge in the droplet is attained. Since the capture of Na^+ ions by PEG involves a transfer of the ions from a water-coordinated configuration into a PEG-coordinated configuration, the critical charge density allows for the cations to be in the vicinity of PEG for a long-lived sodiation of PEG to occur. (ii) It has been discussed in the literature [51] that it is energetically favorable for a Na^+ ion to be solvated in the interior of a cluster or a bulk solution. Close to the Rayleigh limit, the fluctuations due to the charge-induced instability may lower the energy barrier for the release of the Na^+ ion and this may assist the transfer of the ion to PEG.

Regarding the rate of the release of the sodiated PEG, it is found that, once PEG is partially released with the first Na^+ ion, the following ejection events happen at a rapid rate, particularly when the size of the droplet is large, whereas it spends more time until the complete detachment of the macromolecule when the size of the droplet is small. The slower release of the charged PEG from a smaller droplet is attributed to the slower evaporation rate of a smaller droplet. In general, the evaporation rate of a droplet is proportional to its surface area; that is, the smaller the droplet, the smaller the evaporation rate. A slower evaporation rate implies that the charge of the droplet is released more slowly, therefore, it will take longer time for PEG to become fully sodiated and released.

Binding of Na^+ **to PEG** One of the questions to be addressed is whether the sodiation of PEG64 occurs in the interior or on the surface of the droplet. To answer this questions, an additional simulation was performed for a PEG30 molecule in a droplet of 10^5 H_2O molecules and 25 Na^+ ions. This droplet is found slightly below the Rayleigh limit (the corresponding critical charge is approximately 28 *e* using the TIP3P bulk surface tension in Ref. [47]). Thus, this droplet size contains almost the maximum number of single ions that can sustain based on the Rayleigh prediction. When PEG30 was initially solvated at the centre of the droplet, it lingered in the interior of the droplet for approximately 8.5 ns. While in the interior of the droplet, it wrapped one Na^+ ion as the cation was in close proximity to the macromolecule, but the cation was released back to the solvent approximately 2.7 ns after the moment of the sodiation. No further sodiation events were observed as PEG30 diffused outwards to the surface of the droplet. PEG64 in an aqueous droplet of 7000 H_2O molecules, however, resided in the interior of the droplet only for approximately 3.0 ns, and sodiations happened only at the surface of the droplet when PEG settled down on the surface [44–46] (see Figure 2.3).

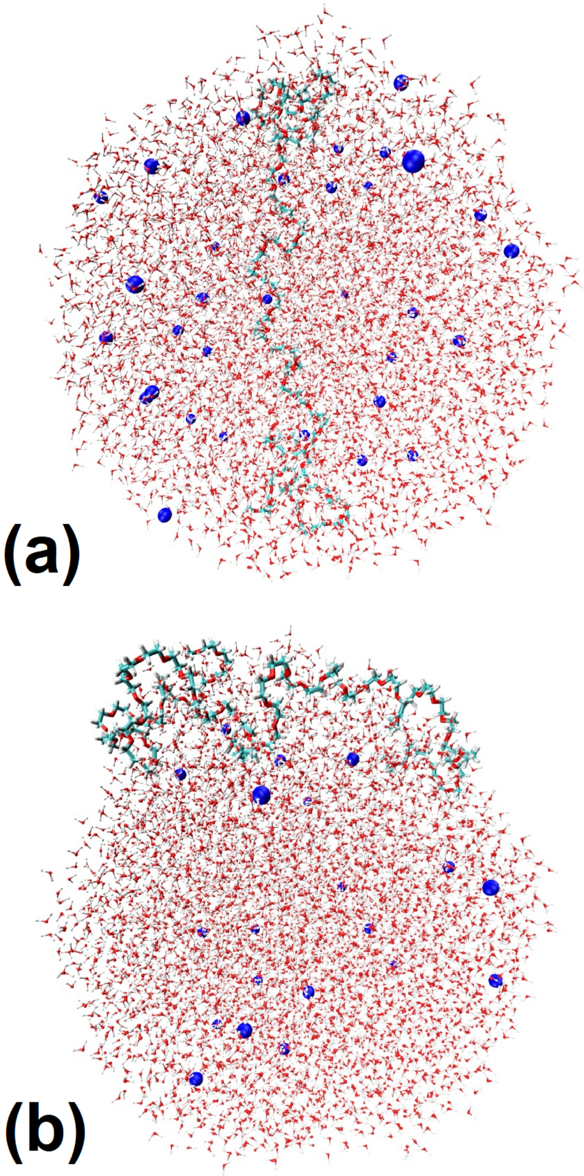


Figure 2.3: Phase separation of PEG in a charged aqueous droplet composed of 32 Na^+ ions and 7000 H_2O molecules.

Therefore, the duration of PEG staying inside the droplet may determine the location of sodiation. In other words, if the macromolecule spends a sufficiently long time inside the droplet, sodiation may happen even before it is found on the surface, provided that the cation is not released back to the solvent while diffusing. The transfer of Na^+ ions from the aqueous solution to PEG involves re-organization of both the solvent and the structure of PEG. This re-organization is facilitated when PEG is on the surface as it has more freedom in its motion since it is not surrounded by solvent.

The current simulations and previous studies [17, 18] indicate that the PEG is located on the surface of a charged droplet. PEG interacts with Na^+ ions via electrostatic attractions due to the presence of negative partial charges on the oxygen sites along its backbone (*i.e.*, ion-dipole interactions). The location of sodiation on PEG is completely random, and the macromolecule can be simultaneously charged with multiple Na^+ ions when it resides on the droplet surface. Once a Na^+ ion is wrapped by a short segment of PEG, it is either released back to the droplet soon (transient sodiation) or it remains wrapped for the rest of the length of the simulation. For transient sodiation, $6-7$ monomers usually form a loop to coordinate the cation. On the other hand, $12-14$ monomers are generally used to form two loops around the cation for long-lived sodiation. Hereafter, this non-transient sodiation will be called highly coordinated sodiation. Transient sodiation always precedes the highly coordinated sodiation, and only the highly coordinated sodiation signals the partial release of the macroion. The Na^+ ion is mostly immobile once wrapped by a short segment of PEG. Figure 2.4 shows the distance between one of the oxygen sites in the loop surrounding Na^+ and the Na^+ ion participating in sodiation as a function of time. This plot clearly distinguishes transient sodiation from highly coordinated sodiation. When a loop is formed around the cation (Figure 2.5), the oxygen atoms point towards the Na^+ ion and carbon atoms point outwards. Therefore, the PEG segment that encloses the Na^+ ion exposes its hydrophobic methylene group towards the aqueous surface, resulting in weaker interactions with the solvent molecules present at the droplet surface. Besides, considering that a Na^+ ion should experience the partial loss of solvation energy as it moves toward the surface, the macromolecule acts as a heterogeneous polymeric solvent, which compensates the unfavorable loss of solvation energy.

The complete removal of excess charge happens when the parent droplet releases a small

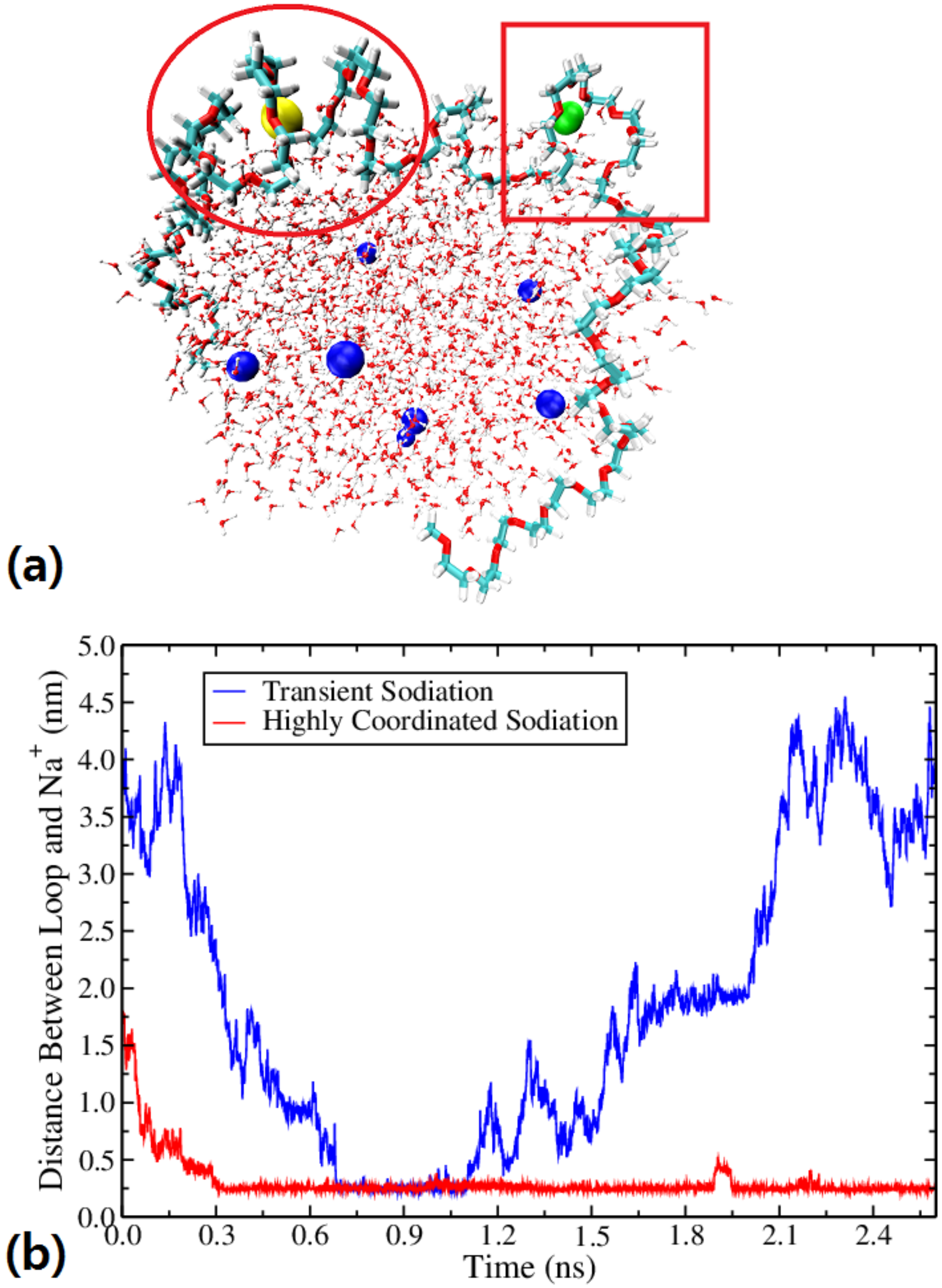


Figure 2.4: (a) Typical snapshot showing transient and highly coordinated sodiation of PEG. The yellow sphere represents a Na$^+$ that is captured by highly coordinated sodiation (the region is encircled) while the green sphere by transient sodiation (the region is indicated by the square). (b) Distance between one of the oxygen sites in the PEG loop and the cation participating in sodiation as a function of time t.

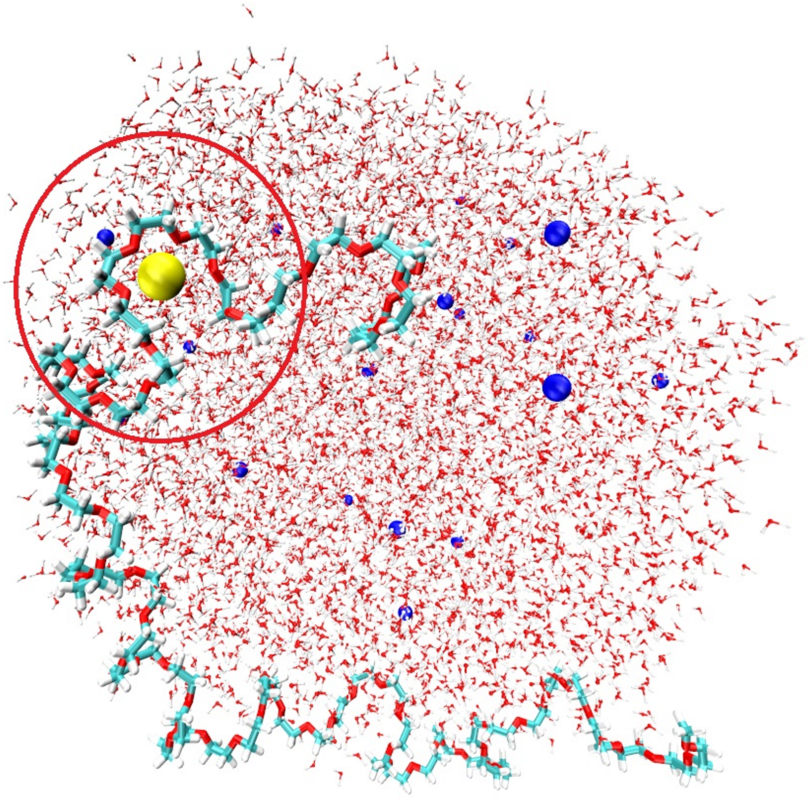


Figure 2.5: Typical snapshot showing the instance of the orientation of the oxygen sites of PEG towards a Na$^+$ ion (red circle).

progeny droplet that contains a single Na^+ ion during Coulomb fissions or when the charged portion of the macroion is ejected. The simulations revealed that (1) in PEG64, the order of sodiation is not always the same as the order of ejection, and that (2) PEG64 is completely detached from the surface of the droplet when it is occupied with four Na^+ ions, regardless of the amount of the initial excess charges. Besides, at the moment of the complete ejection, the Na^+ ions are not uniformly distributed along the macromolecule.

It is worth commenting on the +4 e (where e is the elementary charge) as the final charge state of PEG64. The maximum charge state of PEG64 is +5 e [40]. This maximum charge state is attained in aqueous droplets at relatively low temperature ($\approx$ 280−273 K), whereas, at moderate and high temperatures, the most probable charge state of PEG64 is +4 e. In contrast, the simulations of PEG54 [18] and PEG96 [17] showed that their maximum charge states, which are 4 Na^+ and 7 Na^+, respectively, are attained in a facile manner at various temperatures in the range of 300−350 K. The different maximum charge states are due to the fact that on PEG64 the five Na^+ ions are distributed on the average as one ion per 12.8 monomers, while the five Na^+ ions in PEG54 and the seven Na^+ ions in PEG94 are distributed as one ion per 13.5−13.7 monomers. In PEG64, the +4 e charge state can be easily attained because of the sufficient length of the macromolecule to accommodate this amount of charge.

2.2.3 Electrostatic Model of Charging-Induced Extrusion Mechanisms

In the previous section, a new chain extrusion mechanism has been identified, in which the release of the macromolecule is coupled to its charging on the droplet surface. Hereafter, this mechanism will be referred to as the charging-induced extrusion mechanism (abbreviated as the CI-EM). The findings are not limited to the particular length of the macromolecule and the type of ions [18].

The release of the sodiated PEG from an aqueous droplet is the result of a series of chemical events. The highly coordinated sodiation of PEG does not occur at any droplet size, but at a droplet size that reaches a critical Na^+ concentration. (For instance, as shown in Figure 2.1 (a), this concentration is approximately that of 20 Na^+ ions in a droplet of 6500 H_2O molecules or higher.) Once one sodiated segment of PEG is released, it remains extended at an equilibrium position. To a first approximation, this position is determined by the balance between the solvation force and the electrostatic repulsion. In the following paragraphs, the equilibrium in the extended segment will be analyzed by using the simulation data of PEG extrusion in combination with analytical modelling, in order to obtain insight into factors that play a significant role in the solvation energy. Although an extruding sodiated PEG from an aqueous droplet is used as an example in the discussion, the reasoning can be applied to other extruding macromolecules as well.

The total energy of the system in this analysis is considered to be the sum of its electrostatic and solvation energies. The analysis employs the method of images with the underlying assumption that the droplet is a spherical conductor.

In a typical run, one observes droplet-chain configurations where a partially extruded PEG64 persists for several hundreds of picoseconds. The first extruded part carries a single Na^+ ion (Figure 2.1). q denotes the charge carried by the extruded portion of PEG64. The parameters of the model are shown in Figure 2.6 (a). The electrostatic energy (W^e) of a spherical conductor of radius R and an extruded charge at distance $r = R + \lambda$ (where λ is the length of the extruded portion of the macroion) from the centre of the conducting sphere is given by the following equation:

$$W^e = \frac{1}{8\pi\varepsilon_0}\left[\frac{(Q+q')Q}{R} + \frac{(Q+q')q}{r} - \frac{q'q}{r-r'}\right] \tag{2.1}$$

where $q' = Rq/r$ and $r' = R^2/r$ are the magnitude and the position of the fictitious "image" charge, respectively. The radius of the sphere (droplet) is computed from the total volume comprising water and the solvated part of the chain.

Figure 2.6 (b) shows plots of the electrostatic energy (solid line) using Equation 2.1 and the total energy (dashed line), which is the sum of the electrostatic energy and the solvation energy as a function of λ. In the plots, $Q = +8\ e$, $q = +1\ e$ and $R = 2.1$ nm, which is the radius of a droplet of 1325 H_2O molecules. In Figure 2.6 (b), the first minimum approaches minus infinity for a spherical conductor, which implies that the barrier is infinite. In Equation 2.1, the singularity in the energy is caused by the third term when $r - r' \approx 0$. This singularity is found in the theoretical expressions of the electrostatic energy of a conductor interacting with a charge, but simulations have shown that it is unphysical for the release of ions from droplets.

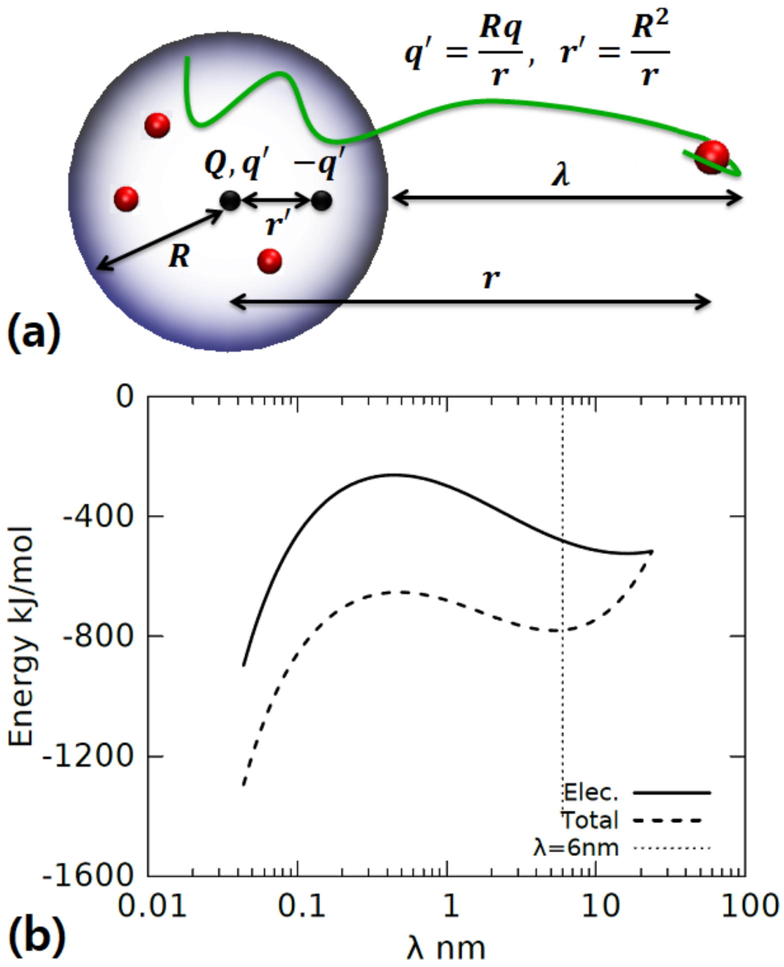


Figure 2.6: (a) Schematic picture that shows the parameters used in the model. The small red spheres and the large blue sphere represent the cations placed randomly and the main droplet, respectively. The green line represents a linear flexible macromolecule. R and Q denote the radius and the net charge of the droplet, respectively. λ is the length of the extruded portion of the macroion. q' is the image charge, while $-q'$ is the counter-charge. $r = R + \lambda$ is the distance between the center of the droplet and the cation attached to the macromolecule, and r' indicates the distance of q' from the center of the droplet. (b) Energy of extrusion of a linear macromolecule from a spherical conductor. The solid line corresponds to the electrostatic contribution given by Equation 2.1 of a single charge removal from a conducting sphere with $Q = +8\ e$, and $R = 2.1$ nm. The dashed line is the sum of the electrostatic and solvation energy contributions for the extrusion of PEG64 from a sodiated aqueous droplet. The dotted vertical line marks the 6 nm extension of the PEG64 from the charged droplet found in simulations. Details of the parameters are presented in the text.

The free energy barrier is finite and this is attributed to the shape fluctuations of the droplets.

For the macromolecule with length L, the solvation energy is expressed as $(L - \lambda)\nu$, where ν is constant solvation energy per unit length. In the simulations, a steady state is observed when the chain was extended 6 nm from the droplet surface. This corresponds to 16 out of 64 monomers leaving the droplet. In Figure 2.6 (b), the dotted vertical line marks the 6 nm extension of the PEG64 from the charged droplet found in simulations. By minimizing the total energy $(W^e + (L - \lambda)\nu)$ at $\lambda = 6$ nm, one finds the value of the solvation parameter $\nu = 16.8$ kJ/nm · mol. Using this value, the total energy (dashed line) shown in Figure 2.6 (b) is estimated. The existence of the barrier shows that the escape of a charge from the conducting sphere is an activated process.

Due to the interaction with an induced charge distribution, in the quasi-static approximation (which is valid due to slow solvent evaporation), the interaction energy of the charge with the droplet next to the surface is dominated by the term $-q^2/16\pi\varepsilon_0\lambda$. As the distance from the droplet increases, the electrostatic (repulsion) energy between the droplet and the charge, $Qq/4\pi\varepsilon_0(R + \lambda)$, overcomes the attraction with the image charge. The next paragraph investigates whether a single constant parameter for the solvation energy can quantitatively explain the chain extrusion from a droplet.

To a first approximation, the electrostatic force exerted on the chain is

$$F = \frac{qQ}{4\pi\varepsilon_0(R + \lambda)^2}. \tag{2.2}$$

Because of the balance established between the forces in the metastable state of the extruded segment, the electrostatic force equals the solvation force $(\partial(L - \lambda)\nu/\partial\lambda = -\nu)$. In Figure 2.7 (a), the dependence of the force on the parameters of the system is evaluated. In the calculations presented in Figure 2.7 (a), a united charge approximation is utilized. In this approximation, it is assumed that all the charge (+1 e, +2 e, +3 e) is located at the end of the chain. The simulations of the ejection of the PEG molecule have demonstrated a non-uniform charge distribution along the chain. However, in the majority of the simulations, all the charges are clustered towards the end of the chain. Hence, to correctly reproduce the electrostatic energy, the united charge approximation is underlying in the model. Within this model, the points shown in Figure 2.7 (a) should fall in a horizontal line (constant value). Instead, the values differ considerably. For instance, for a droplet with 1300 water molecules (the leftmost group of points), the values of the electrostatic force vary by an order of magnitude. Therefore, this model may not capture the major components in the total energy of the system. In the next step of the analysis, it is hypothesized that there is an additional term in the total energy, which is proportional to the electric field on the surface of the droplet.

The simulation findings led to this additional term. The simulations show that the macromolecule lies on the surface of the droplet. It is expected that the local structure of the water on the surface affects the solvation of PEG64 residing on the droplet surface. Therefore, one may anticipate that the local structure of the solvent on the surface will depend on the magnitude of the electric field

$$E = \frac{Q}{4\pi\varepsilon_0 R^2}. \tag{2.3}$$

This assumption was tested by using the linear least square fit of the following model:

$$-\nu + a\frac{Q}{4\pi\varepsilon_0 R^2} = \frac{qQ}{4\pi\varepsilon_0 (R+\lambda)^2} \tag{2.4}$$

where ν and a are the fitting parameters. The parameter ν gives the value for the solvation energy per unit length of the chain. Figure 2.7 shows the results of the fitting in the modified set of parameters $(Q, Q'(l, R, Q, q))$ where

$$Q'(l, R, Q, q) = \frac{R^2}{a}\left[4\pi\varepsilon_0 \nu + \frac{qQ}{(R+\lambda)^2}\right] \tag{2.5}$$

Using R statistical analysis software [52], the following values are obtained for the parameters $\nu = 0.46 \pm 0.03$ and $a = -0.19 \pm 0.02$. The analysis shows that there is a correlation between the electric field on the surface of the droplet and the electrostatic force between the charged extruded segment of the chain and the charged parent droplet. This indicates that the value of the electric field on the surface is potentially connected to the solvation energy. Here, the field is the one prevailing after an equilibrium is established between solvent evaporation and charge evaporation. The change in the droplet size between PEG charging events is very small. The dynamics of the order parameter that corresponds to the degree of chain extrusion is the

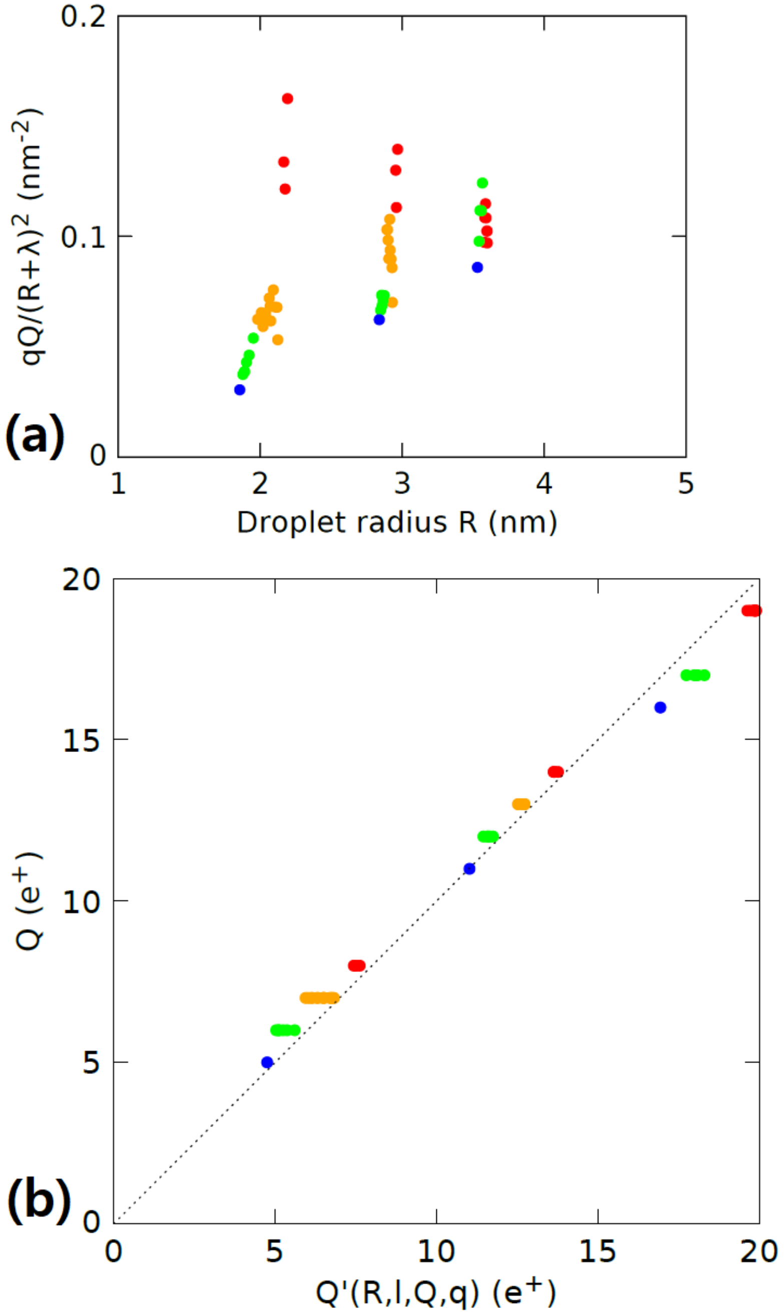


Figure 2.7: (a) Droplet-chain electrostatic force in a united charge approximation explained in the text. Points are colour-coded to indicate the charge on the chain. Chain charges q are +1 e (red), +2 e (orange), +3 e (green), and +4 e (blue). (b) Fit of the data from (a) using a correction to the solvation constant proportional to the magnitude of the electric field on the droplet surface. The points follow the same colour scheme as in (a).

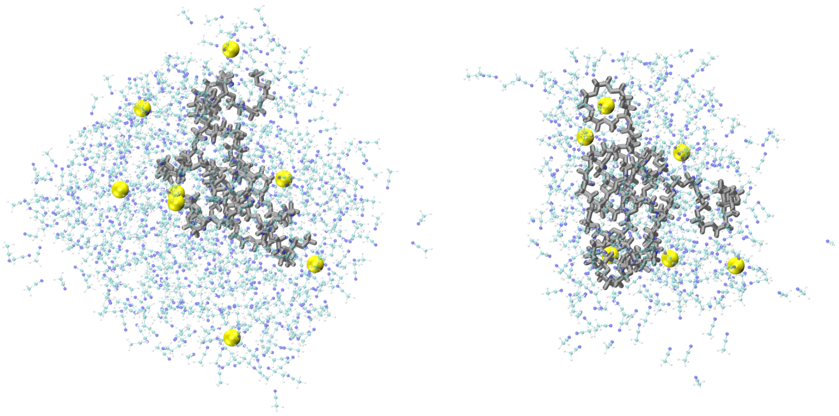
slowest process when compared with the other molecular processes in the system, such as chain conformation dynamics and diffusion of free ions and water molecules. Hence, the solvation force and the electric field can be considered those of an equilibrated system.

2.2.4 Macromolecules in the Interior of an Acetonitrile Droplet

In contrast to the case of water droplets, PEG64 is confined in the interior of a MeCN droplet. The presence of simple charge carriers such as Na^+ ions does not affect the preferential solvation of PEG at the centre of the droplet relative to the neutral MeCN droplet. Consequently, PEG64 is charged with Na^+ ions inside the droplet. This implies that the macromolecule in the organic solvent does not follow the CI-EM. Since the macromolecule remains in the interior of the droplet, it is natural to pose the following question: "Would the ionization process of the macromolecule be the same as the CRM?" This question will be examined in this section.

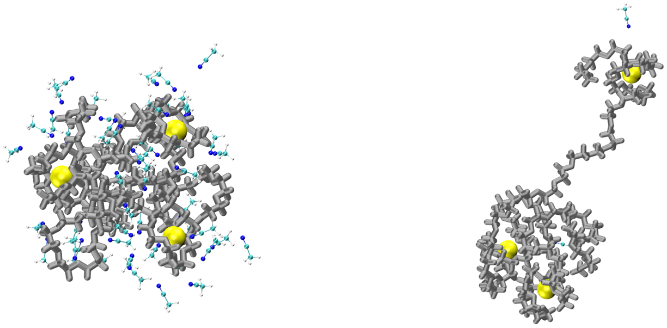
Typical snapshots of the charging of PEG in the interior of an MeCN droplet are shown in Figure 2.8. Unlike the sequential partial releases of the macroion found in a charged aqueous droplet, the MeCN droplet reduces excess charge only through the releases of solvated ions. Hence, the macroion emerges from the parent droplet by solvent drying-out.

In the simulations, the final charge state of the macromolecule is found to be two or four elementary charges. In particular, in MeCN droplets, the initial concentration of the cations and the temperature of the system are significant factors that determine the final charge state of the macromolecule. For instance, at $T = 280$ K, PEG64 obtains a higher charge state (*i.e.*, $+3\ e$) when the droplet initially contains more cations (*i.e.*, $+8-12\ e$). In droplets that start with the same number of the cations, different temperature leads to different final charge states. As the temperature increases, the solvent evaporation rate also increases, thereby accelerating both solvent and ion evaporations. In MeCN droplets, the maximum charge state that PEG64 can sustain is $+4\ e$, which has been obtained at a low temperature ($T = 250$ K). The higher charge state of PEG at the lower temperature may be because of three reasons: (i) the cations have a sufficient amount of time to interact with the macromolecule as solvent evaporation slows down, (ii) the macromolecule does not become compact rapidly as its conformational dimension is strongly correlated to the size of the droplet (see Figure 2.9 (a)), and (iii) the surface tension of solvent generally decreases as temperature increases.



(a) $N_{MeCN} = 747$
$N_{Na^+} = 8$

(b) $N_{MeCN} = 394$
$N_{Na^+} = 6$



(c) $N_{MeCN} = 48$
$N_{Na^+} = 3$

(d) $N_{MeCN} = 2$
$N_{Na^+} = 3$

Figure 2.8: Typical snapshots of PEG64 sodiation and release from a charged MeCN droplet. The Na^+ ions are represented by yellow spheres. PEG64 is coloured in gray. For the purpose of better visualization, the molecular details of PEG64 are not shown. The numbers of MeCN molecules and Na^+ ions remaining in the parent droplet are indicated. The duration of the process shown is approximately 10 ns at $T = 280$ K.

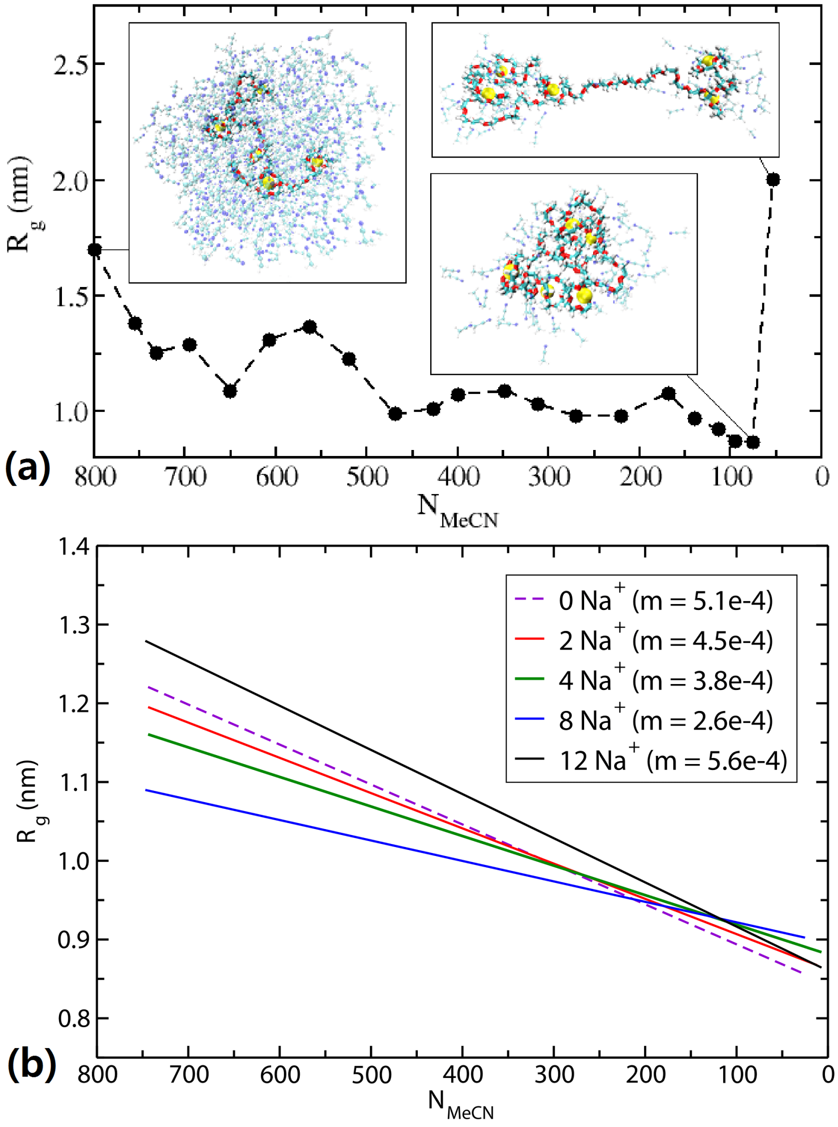


Figure 2.9: (a) R_g of PEG64-5Na$^+$ as a function of the number of remaining MeCN molecules (N_{MeCN}) in an evaporating droplet. The Na$^+$ ions coordinated by PEG64 are represented by yellow spheres. The entire evaporation process after relaxation occurs in approximately 7 ns at $T = 280$ K. (b) Declining trends of R_g of PEG64 in a MeCN droplet with different initial charges. The lines were obtained by fitting the points into a linear function. m is the slope of the lines.

The difference in the final charge state from the case of water droplets is attributed to the different mechanisms of sodiation and ejection of the macromolecule. PEG64 on the surface of an aqueous droplet has an extended conformation that maximizes its access to the ions. On the other end, PEG64 in a charged MeCN droplet becomes more compact as the droplet shrinks (see Figure 2.9 (b)), exposing less for the interactions with the ions. In MeCN, PEG is sodiated only when the cations diffuse close to the macromolecule while the parent droplet loses its mass and charge via evaporation. The charging of PEG in MeCN is therefore different from that predicted by the CRM. The CRM states that during droplet evaporation, single charges are released from the droplet by Rayleigh's mechanism, and when the size of the droplet shrinks to the size close to that of the macroion, complete desolvation causes the remaining charges to collapse onto the surface of the macroion [4, 53]. However, we do not observe the adhesion of the ions onto compact PEG at the latest stage of desolvation; that is, ions can bind to PEG during the evolution of the droplet while solvated ions are released from the parent droplet. Therefore, the comparison between the charging mechanisms of PEG in aqueous and MeCN droplets reveals that the solvent plays a pivotal role in the charging mechanism and the manner in which PEG64 emerges from a droplet. In other words, the role of solvent in the final charge state of a macromolecule is not so simple as the argument proposed from ESI experiments, which states that the final charge state is lower in MeCN simply because of its lower surface tension [14]. The in-depth discussion on this topic is provided in Section 2.2.6.

The conformational changes of the flexible macromolecule residing in the interior of the droplet are significant as the droplet changes its size. To demonstrate this effect, we used the extreme case of PEG64 in its maximum charge state (PEG64-5Na$^+$). In the absence of solvent, PEG64-5Na$^+$ has a linear conformation [40], whereas in the presence of solvent, the macroion is compact. A typical plot of the R_g evolution of PEG64-5Na$^+$ *vs.* remaining MeCN molecules in an evaporating droplet is shown in Figure 2.9 (a). In the same figure, typical snapshots of the change in the conformation of the sodiated PEG64 during solvent evaporation are also included. These trends are common, regardless of the amount of initial charges added into the droplet (see Figure 2.9 (b)).

The general behaviour that is found by examining the stability of the sodiated PEG64 in MeCN droplets is that the macroion stays compact up to the late stages of solvent evaporation.

A single solvation shell is sufficient to maintain a highly charged linear macroion compact. It is at the late stage of evaporation that it partially extends when a small amount of solvent is still present. Depending on the charge state of the macroion, the sodiated PEG bare of solvent will attain a conformation that may be linear or partially unwound [17].

The stability of the highly charged compact PEG64 in small droplets can be explained by the Rayleigh limit. At a first glance, one would expect the Rayleigh limit to not hold for a small droplet that is mainly composed of the charged macromolecule. However, it was found that the partial extension of the macromolecule occurs close to the theoretical predictions of the Rayleigh limit, as shown in Table 2.2. In these estimates, the size of the macromolecule was considered since it is a substantial part of the droplet. Here, it is approximated that each PEG monomer is equivalent to one MeCN molecule, and therefore, the experimental value of the surface tension for MeCN (29 mN/m [54]) was used. This approximation is valid because the surface tension coefficient of MeCN is very close to the one used for the apparent surface energy of PEG globules [20]. Also, Consta and Chung [17] characterized the conformational changes of PEG-$(\mathrm{Na}^+)_n$ ions by using MD simulations, and they observed $Q_c \propto \sqrt{N}$ scaling behavior for the charged PEG (where Q_c is the critical charge and N is the degree of polymerization) as predicted by the experimental findings of Ude *et al.* [20] and the Rayleigh criterion.

It is worth remarking that the relationship between the gas phase conformations of polymer ions and their charge states has been an important question to be addressed in mass spectrometry. For example, Criado-Hidalgo [55] has produced predominantly globular ions of large PEG chains with moderate charge states in negative ESI in order to facilitate ion identification in ion mobility separation coupled to mass spectrometry (IMS-MS). For the ions of small PEG chains (<14 kDa) with charge state (+1 to +5), Ude *et al.* [20] formulated an approximate criterion for the critical mass (m^*) of PEG below which PEG globules lose stability at a Rayleigh-like limit: $m^*(z) \sim 500z^2$ where z is the charge state that values up to 5. This predicts that PEG64 ions ($\approx$ 2.8 kDa) adopts globular geometry only when it bears up to two Na^+ ions, which is consistent with our findings; we found that the conformational shapes of PEG64-$z\mathrm{Na}^+$ are fully stretched when $z = 4, 5$, partially stretched when $z = 3$, and compact when $z = 1, 2$.

Table 2.2: Charge states and droplet sizes of MeCN droplets. The first column shows the number of solvent molecules in a droplet. In the estimate, we assumed that the 64 monomers of PEG are equivalent in size to 64 MeCN molecules. The sum of the solvent molecules and the equivalent PEG64 are shown in parenthesis. The second column shows the simulation-determined charge that a droplet can hold right before the release of charge. The third column shows the Rayleigh limit (R.L.) estimate of the charge that a droplet can hold.

N_{sol} (with PEG64)	Charge State of PEG64 (e)	R.L. Charge (e)
57 (121)	+5	+4.0
20 (84)	+4	+3.3
5 (69)	+3	+3.0

Table 2.3: Radius of the first solvation shell (r) for each solvent and coordination number (CN) of solvent molecules in the first solvation shell to a Na^+ ion for each solvent (H_2O and MeCN) in different solvent compositions (χ_{MeCN}).

χ_{MeCN}	r_{MeCN} (nm)	$r_{\mathrm{H_2O}}$ (nm)	CN_{MeCN}	$CN_{\mathrm{H_2O}}$
1.0	0.300	—	5.89	—
0.8	0.301	0.274	1.24	2.70
0.5	0.302	0.272	0.78	4.81
0.2	0.305	0.268	0.24	4.84
0	—	0.266	—	5.10

2.2.5 Macromolecules in a Water-Acetonitrile Droplet

PEG shows distinct behaviors of solvation and sodiation in pure water and MeCN droplets. Then, an interesting question may arise: what would happen if it is in the mixture of those two solvents? Solvation properties of a water-MeCN binary mixture has been widely studied. Particularly, it has been revealed that micro-heterogeneity in a binary solvent mixture may be displayed as the direct consequence of specific solvent-solvent interactions (*e.g.*, water and some organic solvents such as MeCN [56–60] and dimethyl sulfoxide [59, 61]). In our simulations of H_2O-MeCN mixtures at various compositions, we confirmed the micro-heterogeneity of the binary mixture in a droplet by observing that the size of the water clusters grows as the aqueous component increases, which is consistent with the findings in Ref. [62]. Also, we found that the liquid/vapour interface of the droplet is more enriched in MeCN molecules, whereas water molecules are located mainly in the interior of the droplet. The water molecules are apt to be clustered with one another, many of which surround free and bound Na^+ ions on PEG. Therefore, the presence of solutes may promote the heterogeneous microstructure of the miscible binary solvent when they show preferential solvation by one solvent over the other. For example, the preferential hydration of Na^+ ions has been observed in bulk water-MeCN mixture [63–65], which is consistent with our simulations as shown in Table 2.3.

The preferential solvation of Na^+ ions in water over MeCN is confirmed by computing the radial distribution and the coordination number of a Na^+ ion present in different solvent composition ratios. The maximum coordination numbers of MeCN and water molecules for a single Na^+ are 6 and 5, respectively, which are quantitatively consistent with other computational and experimental studies [66–70]. As shown in Table 2.3, by introducing water molecules into a pure MeCN droplet, the coordination number of the MeCN molecules to each Na^+ ion decreases, and the radius of the first solvation shell by MeCN increases. That is, when χ_{MeCN} = 0.8, the first solvation radius of MeCN decreases by 1.01% and the coordination number drops by 78.8% compared to $\chi_{\mathrm{MeCN}} = 1.0$ (pure MeCN). Moreover, when $\chi_{\mathrm{MeCN}} = 0.2$, the first solvation radius of water decreases by 0.75% and the coordination number decreases only by 5.1% compared to $\chi_{\mathrm{MeCN}} = 0$ (pure water). Vaden and Lisy [71] used infrared predissociation spectroscopy and MP2 calculations to investigate competing noncovalent interactions in ion-

MeCN-water clusters in a gas phase, and they found that the ion-dipole interactions weakens the ion-water electrostatic interactions in the immediate molecular environment of the ion as the large dipole moment of the MeCN is completely aligned with the electric field of the ion. Despite the small change, our results are qualitatively consistent with their results. In both cases, the first solvation shell is interrupted by the introduction of other solvent, yet the degree of the interruption is greater when water is added to MeCN than the reverse case. Therefore, Na^+ ions prefer the solvation by water over MeCN molecules.

Interestingly, we found that micro-heterogeneous solvation, enhanced by the presence of the cations, in combination with differential evaporation rates of the two solvents leads to the formation of the inner aqueous core at the point when the aqueous component becomes dominant. Figure 2.11 shows that the water loss is in a linear trend, whereas the MeCN follows an exponential decay. The order of solvent depletion by evaporation, therefore, is clearly seen: MeCN disappears at a faster rate. The insets show water clusters moving closer to one another and aggregate as the droplet shrinks in size due to solvent evaporation. The formation of an aqueous core is critical in determining the final charge state of the macromolecule in these droplets of binary mixtures.

The formation of the water core is evidenced by the change in the chain dimension of PEG64 as solvent evaporates. In other words, the impact of solvent composition is reflected in the conformations of the macromolecule. As shown in the previous section, when a flexible linear macromolecule (even with a maximum charge state) is located inside a droplet (as for PEG64 in an MeCN droplet), the chain dimension becomes smaller as the size of the droplet decreases by solvent evaporation. When the aqueous component dominates (*i.e.*, $\chi_{\mathrm{MeCN}} = 0.2$), as shown in Figure 2.10, the macromolecule is confined in the interior of the mixture droplet, and simultaneously confined on the surface of the aqueous core. This "sandwiched" solvation at the MeCN/water boundary due to early phase separation of the binary solvent mixture is a direct result of the solvent-solvent and solvent-polymer interactions in water and MeCN. However, the binary solvent mixture has differential evaporating rates for its components (see Figure 2.11), and thus, the conformational dimension of PEG64 is not affected by the depletion of MeCN. The sudden jump at the late stage arises from the partial extrusion of the macroion from the main droplet. On the contrary, when the MeCN component dominates (*i.e.*, $\chi_{\mathrm{MeCN}} =$

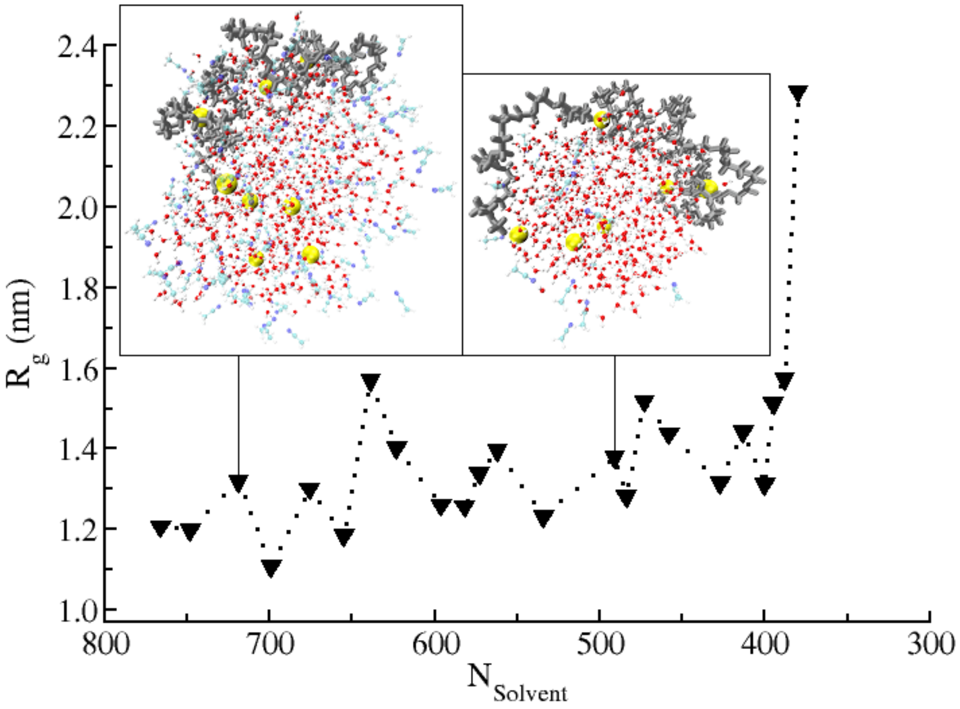


Figure 2.10: Radius of gyration of PEG64 (R_g) *vs*. total number of solvent molecules present in the parent droplet (N_{solvent}) when $\chi_{\text{MeCN}} = 0.2$ at $T = 300$ K. In snapshots, PEG64 is coloured in gray for the purpose of better visualization.

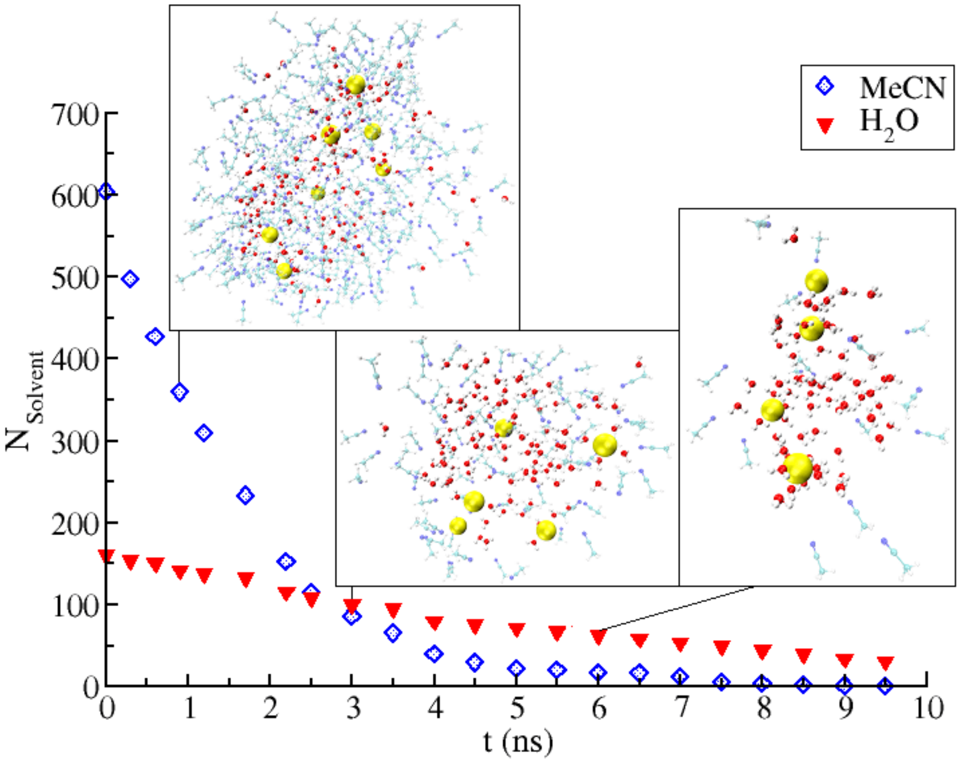


Figure 2.11: The number of solvent molecules remaining in the main droplet (N_{solvent}) as a function of time (t) at $T = 280$ K. The main droplet initially has $\chi_{\text{MeCN}} = 0.8$. PEG64 is hidden in the snapshots to avoid visual confusions.

0.8), the macromolecule is confined in the interior of the mixture droplet. However, because of the lack of the aqueous core at the initial stage, the degree of extension of PEG64 is largely affected by the droplet shrinkage due to the higher evaporation rate of the MeCN component, until the water component dominates by forming the aqueous core. As shown in Figure 2.12 (a), there are two distinct regimes. The blue line represents the MeCN regime where the conformational size of the macromolecule is decreasing according to the solvent (mainly MeCN) evaporation. On the other end, the red line represents the water regime where the chain dimension is kept almost the same due to the predominant presence of the water component along with its slower evaporating rate. When neither component predominates in number (*i.e.*, χ_{MeCN} = 0.5), the two distinct regimes are still found as in $\chi_{\mathrm{MeCN}} = 0.8$. Yet, the range of the MeCN regime is shorter simply because a smaller amount of MeCN molecules are present (Figure 2.12 (b)).

As described in the previous section, PEG64 in charged single-component solvent droplets undergoes distinct solvation, sodiation, and release mechanisms, relying on the location of macromolecular solvation determined by PEG-solvent interactions. The final charge state of the macromolecule in the droplets of the binary mixture, therefore, may be sensitive to the solvent composition of the droplet. However, we found that the final charge state and the manner via which PEG emerges from the droplets can be identical to the case found in pure aqueous droplets. This is obvious in the regime of a higher aqueous component: the charge reduction at the critical limit is dominated by sequential partial releases of the macroion at the late stage. In the regime of a lower mole fraction of MeCN, a series of Coulomb fissions occur at the early stage as a predominant mechanism to eliminate excess charges from the parent droplet, and PEG64 is extruded in a sequential manner when only the small water droplet is left. As shown in Table 2.4, the number of charge-reducing events (Coulomb fissions *vs.* macromolecular partial releases) before the complete detachment of PEG64 from the droplet is different, depending on the solvent compositions. Yet, the number of the cations remaining in the aqueous core is larger than or equal to four in all composition ratios, owing to the preferential hydration of the ions. In other words, even at the small mole fraction of water (*i.e.*, $\chi_{\mathrm{MeCN}} = 0.8$), PEG64 can be liberated from the parent droplet through sequential partial releases as in a charged pure aqueous droplet, leading to the final charge state of +4 *e*. This is

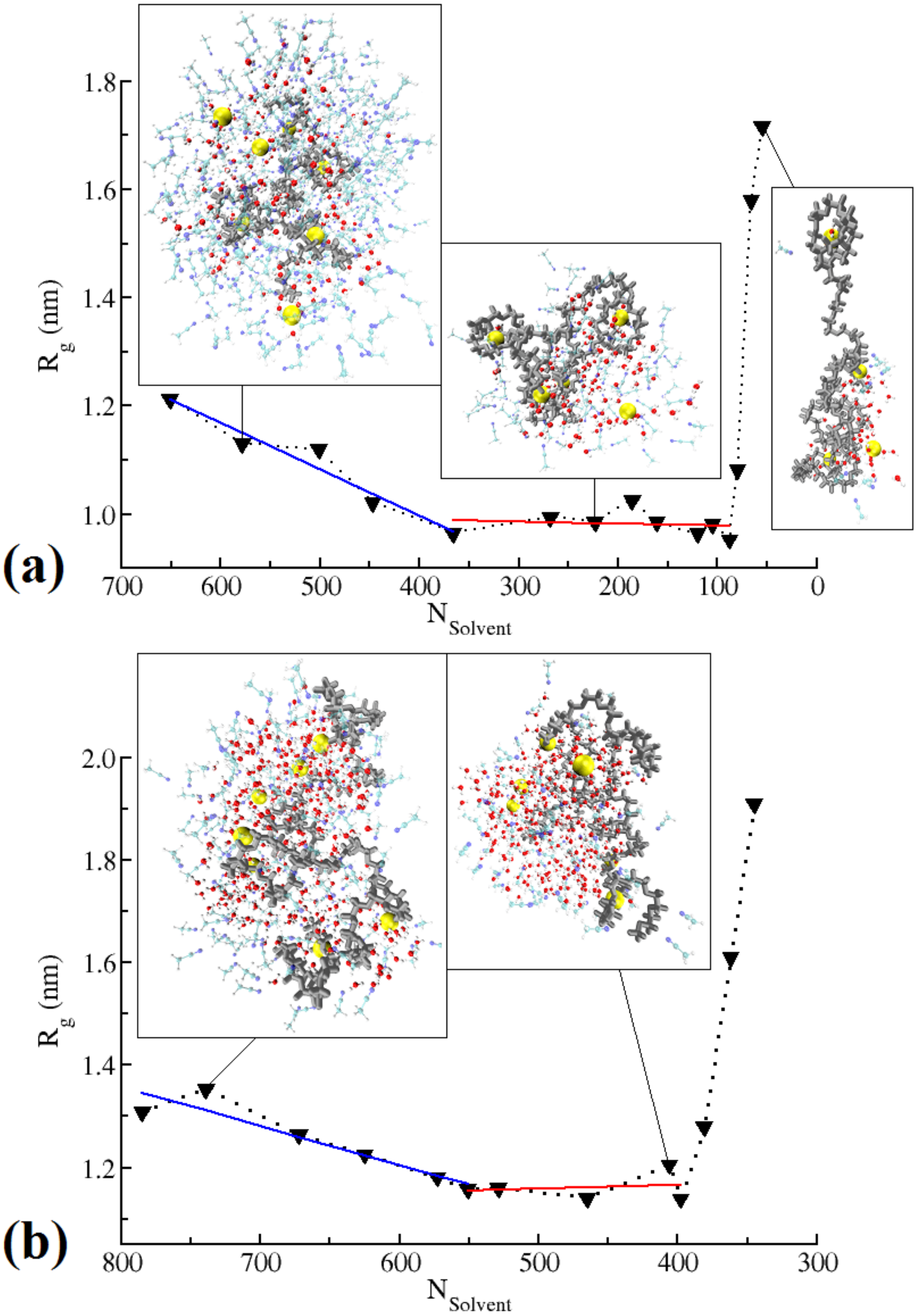


Figure 2.12: Radius of gyration of PEG64 (R_g) *vs*. total number of solvent molecules present in the parent droplet (N_{solvent}) when (a) $\chi_{\text{MeCN}} = 0.8$ and (b) $\chi_{\text{MeCN}} = 0.5$ at $T = 300$ K. Blue and red solid lines represent the MeCN and water regimes, respectively, affecting the chain dimension of PEG64. In snapshots, PEG64 is coloured in gray on the purpose of better visualization.

Table 2.4: Number of Coulomb fissions (N_{CF}), number of free Na^+ ions before the first partial release of the macroion (N_{Na^+}), and final charge state of PEG64 in different solvent compositions (χ_{MeCN})

χ_{MeCN}	N_{CF}	N_{Na^+}	Final Charge State
0	0	8	PEG64-4Na^+
0.2	2	6	PEG64-4Na^+
0.5	2	6	PEG64-4Na^+
0.8	4	5	PEG64-4Na^+
1.0	5	3	PEG64-3Na^+

attributed to the formation of the inner water core in combination with preferential hydration of the cations.

2.2.6 Comparisons with Experimental Observations

The role of surface tension in the determination of the final charge state of a macromolecule has been questioned in ESI experiments [72–78]. It has been found that higher charge states can be observed when the surface tension of the liquid raises, as the droplets can hold more charges which would be otherwise lost during droplet disintegrations at the early stage of the droplet evolution [13]. Also Iavarone and Williams [14] demonstrated that in the absence of other factors, the surface tension of the droplet in the late desolvation process is a critical factor for overall analyte charge by investigating the extent of charging of different analytes, including PEGs, with ESI-MS. They used methanol and water as single-component solutions and added *m*-nitrobenzyl alcohol (*m*-NBA) to alter the surface tension of solvent. They found a strong correlation between the surface tension of solvent and the final charge state of the macroion. Even though there is still the ongoing debate on the role of surface tension (and its relevance to protein supercharging [78]), it seems that their experimental observations are consistent with the simulations of charged droplets of pristine water and acetonitrile in this study. According to the Rayleigh mechanism, it is expected that the addition of an organic solvent with lower surface tension and lower volatility than water should lead to a lower charge distribution of analytes. However, our findings reveal that in addition to the surface tension of solvent, more factors come into play in the charging mechanism of the macromolecule. The difference in the charge state of PEG is attributed mainly to the combination of the location of PEG in a droplet (surface *vs.* interior) and how rapidly the solvent evaporates. We found that PEG64 ends up with four Na^+ ions when solvated in water droplets, whereas it often takes less than four charges when solvated in acetonitrile droplets. This is because (i) PEG64 is on the surface of the water droplet, adopting more extended conformations, (ii) the location of solvation of the macromolecule alters the manner that PEG is charged and released, and (iii) more of the free Na^+ ions floating in acetonitrile droplets escape the parent droplet by Coulomb fissions due to its faster evaporation rate. Interestingly, however, experimental results countering the significance of surface tension have been provided [73–77]. Also we found that the presence of MeCN in an aqueous droplet may not have a critical impact on the release mechanism and the final charge state of PEG64, at least for the three composition ratios we examined in this article (*i.e.*, χ_{MeCN} = 0.2, 0.5, and 0.8), although the addition of MeCN into an aqueous solution is known to decrease the surface tension of the resulting mixture [79]. As PEG shows distinctive ejection mechanisms with varying final charge states based on the type of solvent, it was strongly expected that the macromolecule would be induced to have intermediate degrees of those properties, depending on the solvent composition ratio. However, our finding is completely different: the introduction of a MeCN component into a H_2O droplet does not change the charging and ejecting mechanisms, and therefore, PEG64 in these binary droplet systems shares the common features of the IEM. We attributed this to micro-heterogeneity, different solvent evaporation rates, and preferential hydration of Na^+ ions.

In order to study the effect of solvent on the maximum charge state and charge state distribution of proteins generated by ESI, Iavarone *et al.* [80] compared electrospray spectra of cytochrome *c* and myoglobin in the solution mixture of 47% water/50% organic solvent/3% acetic acid, where the organic component was methanol, MeCN, or isopropanol. They claimed that the charge state distributions of the two proteins shift to lower charge with a noticeable reduction in the abundance of the maximum charge state. Similarly, Hopper *et al.* [81] placed a small reservoir of MeCN into the atmospheric pressure region of an electrospray source, and they found that the average charge states of proteins (trypsin-benzamidine protein-inhibitor complex and human transthyretin) sprayed from aqueous buffers were lessened by exposing electrospray droplets to the neutral solvent vapour. According to their studies, the effect of MeCN molecules on the final charge state of a macromolecule is not negligible, whether they are present inside or outside the droplet, as they result in the decrease in the charge states of the macromolecular ion. This is apparently not in accord with our results, because the cationization of PEGs is typically dominated by metal cation adduction [14, 19], whereas proteins have more complicated architectures and both covalent and noncovalent interactions that PEGs do not have. For instance, it has been commonly known that the charge state distribution of a protein ion is influenced significantly by its molecular conformation and acid-base chemistry both in solution and in the gas phase [80]. Denatured, elongated protein structures are prone to have higher charge states than their folded tertiary structures, due to higher affordability of the electrostatic repulsions between neighboring charges and higher accessibility of basic residues in the protein [80]. Also the charge transfer between a protein and solvent molecules may involve proton transfers in addition to the adduction of a charged species. These aspects of proteins make it difficult to extend the findings in this study to the case of protein ions generated by ESI.

2.3 Extrusion Mechanisms of Proteins from Droplets and "Droplets from Macroions"

2.3.1 Modelling and Computational Methods

An aqueous droplet that involves each of the following proteins was simulated: (a) the transmembrane domain of Bcl-2 member Harakiri (RCSB PDB code [82, 83] 2L5B [84]), (b) the heat shock protein 12 (Hsp12) from Saccharomyces cerevisiae (RCSB PDB code 2LJL [85]), and (c) myoglobin (RCSB PDB code 1MBN [86]).

All the proteins and ions were modelled using the CHARMM General Force Field [87, 88]. The H_2O molecules were modelled using the TIP3P model [89]. Constant-energy and constant-energy MD simulations were performed by using the NAMD package version 2.9 [90]. All MD trajectories were integrated by using the velocity-Verlet algorithm with a time step of 1.0 fs. In order to perform equilibrium MD simulations, every system was confined in a spherical cavity of radius 50 nm, which was generated by applying the spherical boundary condition. The cavity was sufficiently large to accommodate the largest droplet shape fluctuations. Equilibrium between the droplet and water vapour was established inside the cavity as some water molecules evaporated and filled the vacant space in the cavity after relaxation. The droplet systems were then thermalized for 1−3 ns by using Langevin dynamics or coupling them to a heat bath of a specific temperature. After equilibration, the thermostat was turned off and the runs were continued using constant-energy MD. Both electrostatic and van der Waals interactions were explicitly treated with a switch scheme. The cutoff distances were set so that every interaction within the main droplet was included in the force calculations.

The surface tension and the density values of the TIP3P water model were obtained from Refs. [47, 48, 91, 92]. VMD version 1.9.2 [38] was used for visualization of the trajectories.

Membrane Protein-Droplet Systems An aqueous droplet consisting initially of 1000 H_2O molecules, six Na^+ ions, and 2L5B with protonation state +2 *e* (maximum charge state) was simulated at two different temperatures, 290−300 K and 320−330 K. After equilibration, the main droplet at room temperature contained 950−970 H_2O molecules. The corresponding Rayleigh limit is approximately 9−10 *e*, and therefore, the system is slightly (≈ 11−20 %) below the Rayleigh limit. At the elevated temperature, the emission of a solvated ion was observed within 1 ns of equilibration.

IDP-Droplet Systems 2LJL with different charge states was solvated in two aqueous droplets of different sizes: (a) $2LJL^{15+}$ in a droplet that was initially composed of 1500 H_2O molecules at 320−340 K, and (b) $2LJL^{15+}$ in a droplet that was initially composed of 3000 H_2O molecules at 310−320 K. The systems were first thermalized at 350 K for (b) and at 320 K for (c).

In (a), a charge-induced unfolding was observed when the system contained 1210−1320 H_2O molecules. The connected droplet was left with approximately 1000 water molecules after 10 ns. The Rayleigh limit of this system is approximated to be 11−12 *e*. Therefore, the system is quite (≈ 25−36%) above the Rayleigh limit. In (b), 2850−2860 water molecules remained in the droplet after equilibration. The Rayleigh limit for this size of the system is 15−16 *e*. Therefore, the system is at the Rayleigh limit.

Denatured Protein-Droplet Systems A multiply protonated 1MBN in water was studied in seven different initial settings: (a) Compact $1MBN^{17+}$ in 1070 H_2O molecules at 340−370 K,

(b) Compact $1MBN^{17+}$ in 2500 H_2O molecules at 310−320 K, (c) Compact $1MBN^{17+}$ in bulk solution with 17 Cl^- ions at 1 atm and 320 K, (d) Compact $1MBN^{36+}$ in 3000 H_2O molecules at 310−320 K, (e) Extended $1MBN^{36+}$ in 6460 H_2O molecules at 310−320 K, and

The charge state of +17 *e* was obtained from the ESI mass spectra found in Ref. [26], and the charge state of +36 *e* was obtained from the maximal protonated state that the protein can have.

System (a) was thermalized at 350−380 K, and all the other systems at 310−320 K. The initial droplet size in all the systems was reduced because of solvent evaporation until an equilibrium between the vapour and the droplet has been built. As systems (a), (d), and (e) were equilibrated, they were found substantially above their Rayleigh limits due to solvent evaporation. The droplet was found in the instability regime. The instability was manifested by charge-induced unfolding of the protein. When charge-induced unfolding occurred, the system was not still at equilibrium. The equilibrium state was reached after the unfolding of the protein. Systems (b), on the other hand, is at the Rayleigh limit.

In system (a), only 530−550 water molecules remained in the connected body of the droplet at the moment of charge-induced unfolding of the protein. The rest of the H_2O molecules were found in the vapour phase within the cavity. The Rayleigh limit of the droplet of this size (accounting the volumes of both $1MBN^{17+}$ and H_2O molecules) is approximately 9−10 *e*. Therefore, the system is found $\approx$ 70−89 % above the Rayleigh limit. The charge-induced unfolding was followed by further evaporation of H_2O molecules from the droplet until the equilibrium between the vapour and the droplet was reached. At the later period of the run, the droplet was composed of $\approx$ 350 H_2O molecules.

In system (b), the droplet contained $\approx$ 2400 H_2O molecules after 1 ns of equilibration. In contrast to the $1MBN^{17+}$ in (a), the protein ion remained compact for 10 ns in (b). Taking into account the volumes of both $1MBN^{17+}$ and H_2O molecules, the Rayleigh limit of this droplet is 16−17 *e*. Therefore, the system is at the Rayleigh limit.

In system (c), $1MBN^{17+}$ in bulk solution was simulated by locating it at the centre of the periodic cell (9 nm × 9 nm × 9 nm) filled with water molecules and 17 chloride (Cl^-) ions. The system was first equilibrated for 1 ns at 1.01325 bar and 320 K by using the Langevin piston Nose-Hoover method for pressure and Langevin dynamics for temperature, and then the barostat was turned off after 0.5 ns. For the production run, the NVT simulation continued for 10 ns from the last time frame of the equilibration run. The van der Waals interaction was truncated smoothly by the switching function with the switch and cutoff distances of 10 Å and 12 Å, respectively, while Particle Mesh Ewald was used to treat electrostatics with the periodic boundary condition.

In system (d), $\approx$ 2900 H_2O molecules remained in the droplet at the moment of charge-induced unfolding of the protein at 320 K. This size of the droplet, accounting the volumes of both $1MBN^{36+}$ and H_2O molecules, corresponds to the Rayleigh limit of 16−17 *e*. Therefore, the system is found $\approx$ 112−125 % above the Rayleigh limit. The number of the remaining solvent molecules in the droplet then decreased to $\approx$ 2750 in the later stage of the run.

In system (e), the droplet was simulated only for 1 ns (equilibration) at 320 K. Because of the lower temperature and the highest charge of the 1MBN, the most of the H_2O molecules ($\approx$ 6177) remained in the connected body of the droplet. The Rayleigh limit is approximately 23−24 *e*, and therefore, the system is $\approx$ 50−57 % above the Rayleigh limit. Within this time, the charge-induced instability was manifested by the splitting of the droplet into two lobes clinging to the termini of the protein.

2.3.2 Solvation of Charged Proteins in Aqueous Droplets

Membrane Proteins In 2L5B, there are only two basic amino acids, arginine-28 and arginine-29, which are protonated to give the net charge +2 *e* of the protein at pH 7 or lower. These amino acids are located in the proximity of the unstructured tail of the protein. Although the protein resides on the droplet surface due to its high degree of hydrophobicity (Figure 2.13 (a)), the droplet lowers its charge state by emission of solvated Na^+ ions and no release of the protein (Figure 2.13 (b)). This is observed in the identical system at the elevated temperature.

In order to understand the reason that a charged hydrophobic protein is not released from an aqueous droplet, the solvation of the PEG64 is compared with that of $2L5B^{2+}$. In the comparison, the following two systems are used: (a) a droplet composed of 950−970 H_2O molecules, six Na^+ ions, and the $2L5B^{2+}$ ion and (b) a droplet composed of 1550−1560 H_2O molecules, nine Na^+ ions, and a PEG64 molecule. In both of the aqueous droplets, the macromolecules

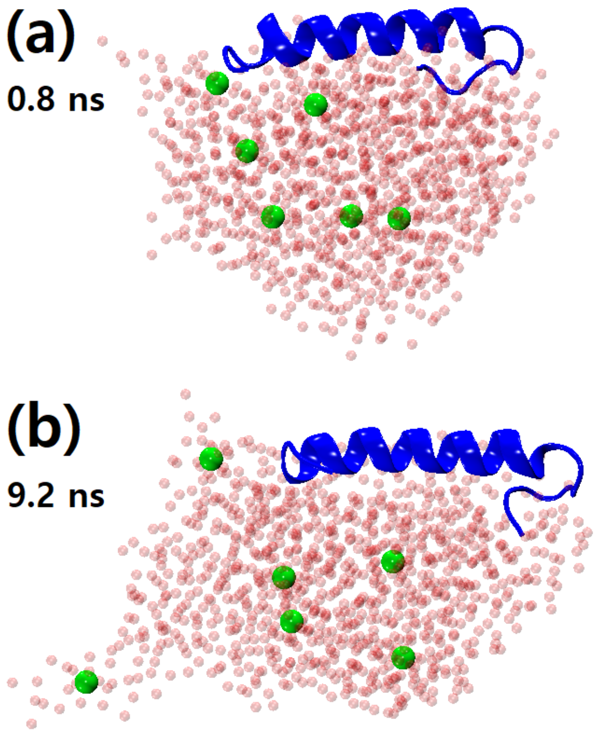


Figure 2.13: Fission of an aqueous droplet that contains the hydrophobic trans-membrane domain 2L5B^{2+}, six Na^{+} ions, and 950-970 H_2O molecules. The snapshots show the droplet system at two different time frames, namely, (a) 0.8 ns and (b) 9.2 ns after the system relaxation. The protein and water molecules are coloured in blue and light red, respectively. The cations are represented by green spheres.

reside on the droplet surface. Two numerical quantities were computed to measure the attachment of water to the macromolecules: the time-average of the number of water molecules in the first few solvation shells ($r \leq 5$Å) per solvent-accessible surface area for each monomer (χ), and the time-average number of hydrogen bonds that form between water molecules and the macromolecules ($\overline{n}_{HB}$).

Figure 2.14 shows the χ values for the 31 residues in 2L5B (left) and those for the 64 monomers in PEG (right). 2L5B exposes one side of the helical structure to the vapour phase (green dots with low χ values) and the other side to the droplet phase (red dots with high χ values). This causes a large difference in χ between $\chi_{max} = 46.2$ and $\chi_{min} = 0.2$. The mean of the χ values of the red dots is $\overline{\chi}_{\mathrm{2L5B}} = 19.0$. On the other hand, the fluctuation of the χ values for the blue dots is attributed to the dynamic formation of loops in PEG [93] that alternate between the droplet interior and the vapour phase as shown in Figure 2.15. The mean of the χ values of the blue dots is $\overline{\chi}_{\mathrm{PEG}} = 14.9$. The quantitative analysis of hydrogen bonding between solvent and the macromolecules reveals that $\overline{n}_{HB,\mathrm{2L5B}} = 33.4 \pm 4.4$ and $\overline{n}_{HB,\mathrm{PEG}} = 7.8 \pm 3.6$ (where the errors were obtained from the standard deviations of the corresponding data sets). Therefore, both quantities indicate that the protein ion has stronger attachment to water than PEG. The different solvation of the membrane protein from that of PEG gives rise to different mechanisms by which they emerge from the droplet. The higher the charge of a membrane protein, the more difficult the detachment from an aqueous droplet would be.

Moreover, although the ionizable residues are present at the unstructured tail of the protein, even partial extrusion of the chain would not take place. This is because these residues are located in the middle of the sequence of non-charged residues (including glycine-27, asparagine-30, and leucine-31), and hence, the extrusion should take all these neighbouring residues to the vapour.

Intrinsically Disordered Proteins To examine the possible extrusion mechanism of an IDP from a water droplet, $\mathrm{2LJL}^{15+}$ is placed in water droplets of two different sizes: (a) 1500 H_2O (above the Rayleigh limit) and (b) 3000 H_2O (at the Rayleigh limit). 2LJL is composed of ≈ 28.4 % (31 out of 109 amino acids) hydrophobic amino acids.

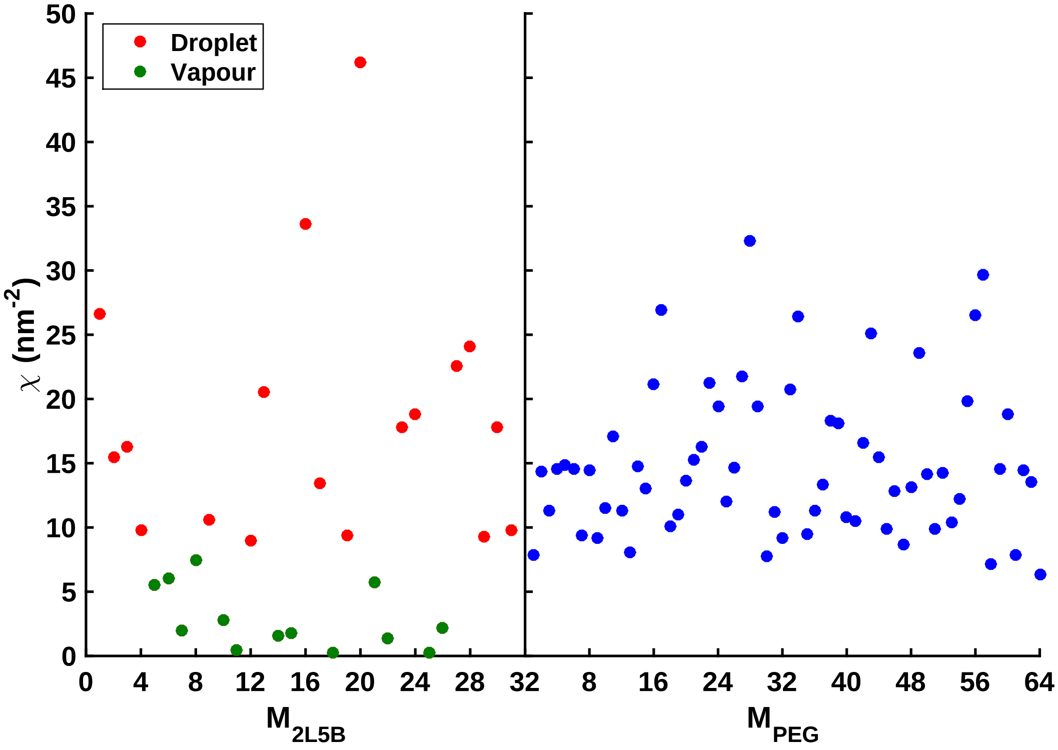


Figure 2.14: χ (see the main text for the definition of the quantity) values of the 31 residues in 2L5B, averaged over 10 ns (left) and the 64 monomers in PEG64, averaged over 1.5 ns (right). M_{2L5B} and M_{PEG} denote the residue number of 2L5B and PEG, respectively. The red dots indicate the residues in direct contact with water, whereas the green ones indicate those with exposure to the vapour. The blue dots are for PEG monomers.

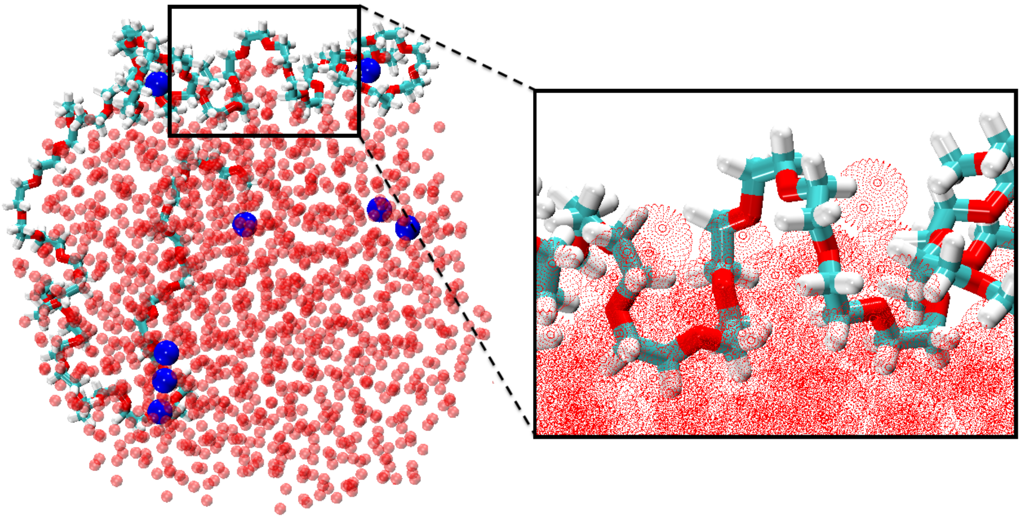


Figure 2.15: Typical snapshot of PEG64 in an aqueous droplet charged with Na^+ ions. The Na^+ ions are presented by the blue sphere, and the water molecules by the red spheres. The inset shows the loops of PEG on the water surface, and thus, its reduced contact with water.

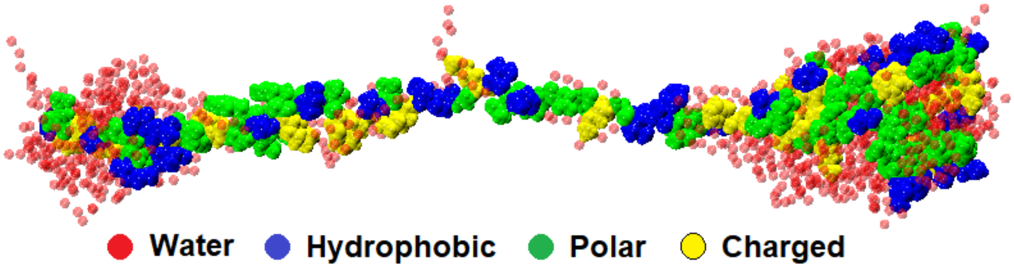


Figure 2.16: Equilibrium state of the intrinsically disordered protein $2LJL^{15+}$ in an aqueous droplet of approximately 1000 H_2O molecules. The nature of the amino acids throughout the chain backbone is shown.

Figure 2.17 shows the evolution of the droplet from a compact transient form to a stable extended conformation in system (a). The droplet above the Rayleigh limit is unstable, and thus, the system undergoes a transient stage where $2LJL^{15+}$ extends rapidly without passing from the intermediates states of gradual extrusion like the charged PEG. As the chain stretches, it also drags the H_2O molecules with it. In a transient state, the elongated $2LJL^{15+}$ is almost uniformly solvated by H_2O molecules (Figure 2.17). As shown in Figure 2.17, hydrophilic and charged amino acids are distributed along the $2LJL^{15+}$ backbone. Instead of solvating the chain almost uniformly, the water molecules are expelled from the central body of the already extended macroion, leaving the central part of the chain with a thin sheath of water. This sparse distribution of water along the middle region of the chain is due to hydrogen bonding and electrostatic interactions between water and the amino acids. Therefore, the water molecules accumulate at the chain termini by forming two “pear-shaped” lobes.

The $2LJL^{15+}$-droplet conformation shown in Figure 2.16 is an equilibrium state under certain conditions of temperature and pressure. For $2LJL^{15+}$ at 350 K, the droplet at one terminus contains 246 H_2O molecules and 23 residues. In this lobe the overall charge is +4 *e*, which is lower than the Rayleigh limit +5 *e*. In the other terminus the droplet contains 685 H_2O and 56 residues. The charge state of the lobe is +7 *e*, while the Rayleigh limit is +8 *e*. This droplet morphology is consistent with the “pearl-necklace” scenario shown in Figure 2.22 (b). The explanation of the two-lobe structure is discussed in the next section, where a highly charged myoglobin shows the same structure as the $2LJL^{15+}$.

When the $2LJL^{15+}$-droplet is close to the Rayleigh limit, neither the “pear”-shaped water lobes nor the chain extrusion is observed. As shown in Figure 2.18, one or two thin rays are developed on the droplet surface. The protein ion is instead packed in the interior of the droplet. The formation of the rays indicates that a transition from the “star”-shaped morphology to the “pear”-shaped morphology will take place as the charge-squared-to-volume ratio of the droplet increases with reduction in the droplet size. The lifetime of the “star”-shaped morphology is determined by the stability of the protein against charge-induced chain stretching and the rate of the stretching if occurs.

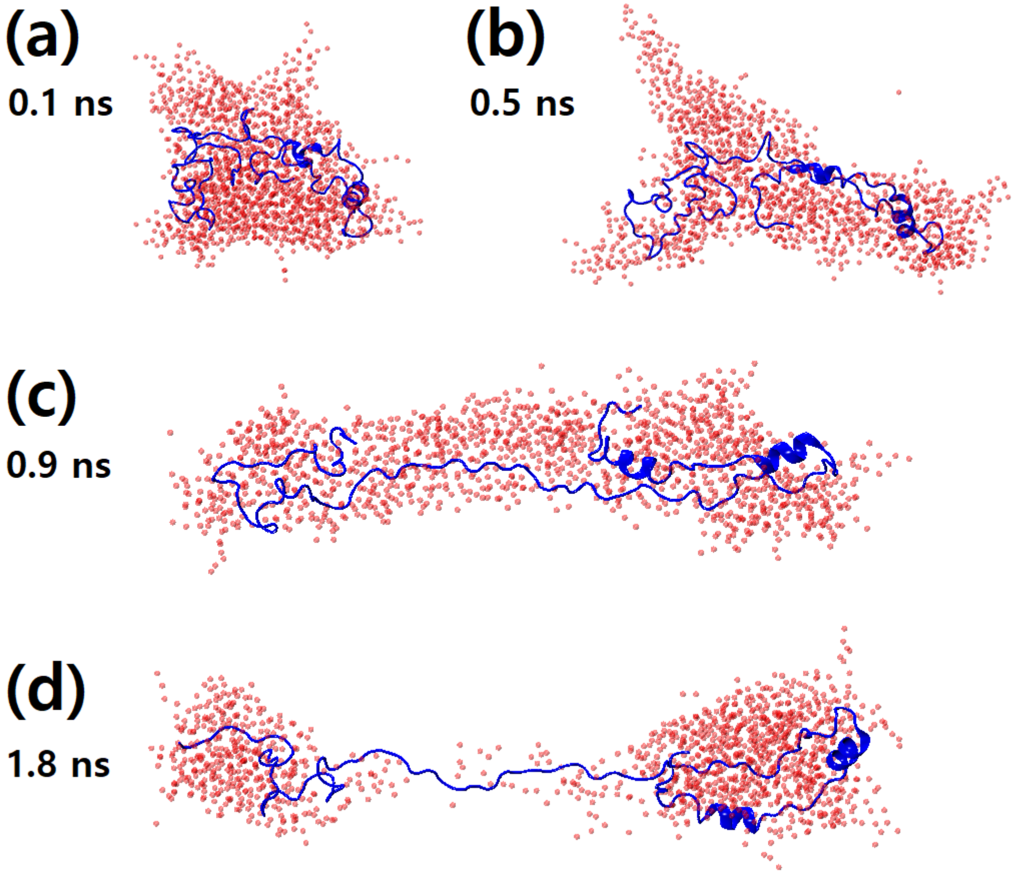


Figure 2.17: Charge-induced stretching of the intrinsically disordered protein 2LJL^{15+} in an aqueous droplet of 1210-1320 H_2O molecules. The snapshots show the time evolution of the droplet system: (a) 0.1 ns, (b) 0.5 ns, (c) 0.9 ns, and (d) 1.8 ns during equilibration. The molecular models are fully atomistic, but some of the atoms are not explicitly displayed in order to facilitate visualization. The protein is coloured in blue, and the oxygen sites of water molecules are represented by transparent red spheres.

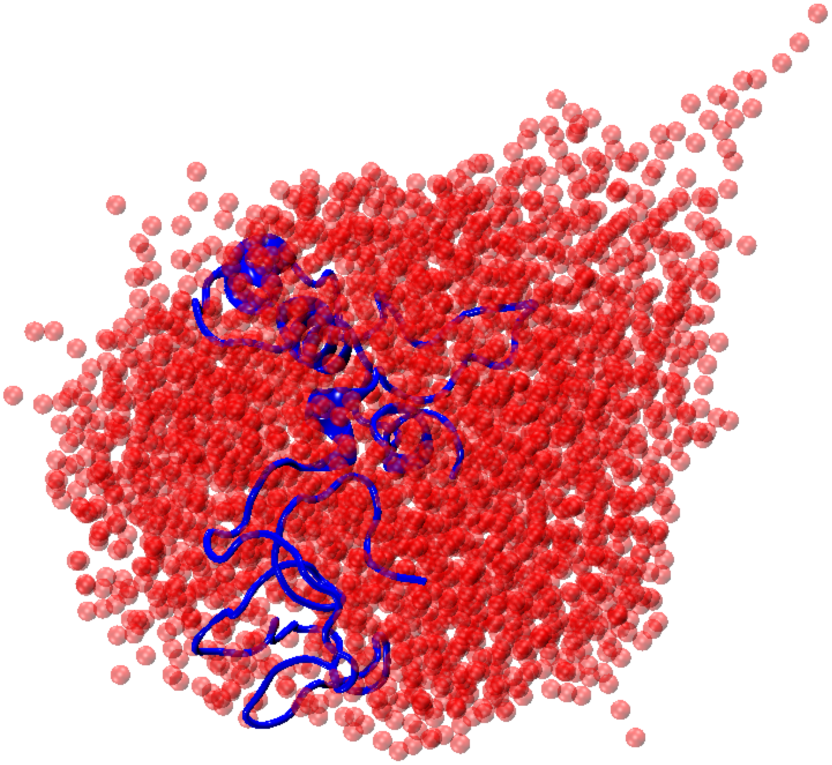


Figure 2.18: $2LJL^{15+}$ in an aqueous droplet of 2850−2860 H_2O molecules. The same colour scheme is used as in Figure 2.17.

Denatured Proteins In this and following paragraphs, the morphologies of aqueous droplets involving $1MBN^{17+}$ and $1MBN^{36+}$ will be examined. 1MBN contains $\approx$ 43.1 % of hydrophobic residues in its total sequence of 153 amino acids. If occurs, $1MBN^{17+}$ and $1MBN^{36+}$ in aqueous droplets show very similar solvation and unfolding mechanisms. $1MBN^{17+}$ in bulk solution and in a droplet below the Rayleigh limit acquires a compact conformation as it is evidenced by the root-mean-square displacement (RMSD) shown in Figure 2.19. Our simulations reveal that $1MBN^{17+}$ remains compact in a aqueous droplet close to the Rayleigh limit. As expected, $1MBN^{36+}$ obtains an extended conformation when it is bare of solvent. In an aqueous droplet, $1MBN^{36+}$, although denatured, obtains a coiled or somewhat compact conformation when the system is found below the Rayleigh limit.

The solvation of $1MBN^{36+}$ in an aqueous droplet above the Rayleigh limit will be described in this paragraph. The same discussion is applied to $1MBN^{17+}$ in the droplet of 530−550 H_2O molecules, by the reason of its resemblance to the evolution of $1MBN^{36+}$-droplet systems (see Figure 2.20). The conformational changes of $1MBN^{36+}$ in aqueous droplets with $\approx$ 6100 and $\approx$ 2900 H_2O molecules were monitored by preparing extended and compact initial conformations of $1MBN^{36+}$, respectively. Both systems are found substantially above the Rayleigh limit.

In a droplet of $\approx$ 6100 H_2O molecules, the water migrates to the termini of an extended $1MBN^{36+}$ forming two lobes (see Figure 2.21 (c)). Figure 2.20 (a) shows the equilibrium state of $1MBN^{36+}$ in a droplet of approximately 6100 H_2O molecules. One terminus accommodates 3050 H_2O molecules and 52 residues with charge $+12\ e$ (the corresponding Rayleigh limit is $+16e$). The lobe at the other end is composed of 3030 H_2O molecules and 61 residues at a charge state $+14\ e$ (the corresponding Rayleigh limit is $+16\ e$). This conformation is consistent with the droplet splitting scenario observed in the $2LJL^{15+}$-droplet system, and a uniformly charged PEG in a methanol droplet [94]. These simulations show that the extended part of the chain emerges because of the stretching of a highly charged protein due to charge-induced instability, not because of chain extrusion as it occurs in the charged PEG from a water droplet (see the first section). As the unfolding occurs, the water of the droplet is dragged along the chain and surrounds it as it was described for $2LJL^{15+}$. $1MBN^{36+}$ in a compact conformation in a droplet of $\approx$ 2900 H_2O molecules also extends and forms aqueous lobes at the termini, as seen in Figure 2.20 (b). The equilibrium state of $1MBN^{17+}$ is shown in Figure 2.21 (b). Since

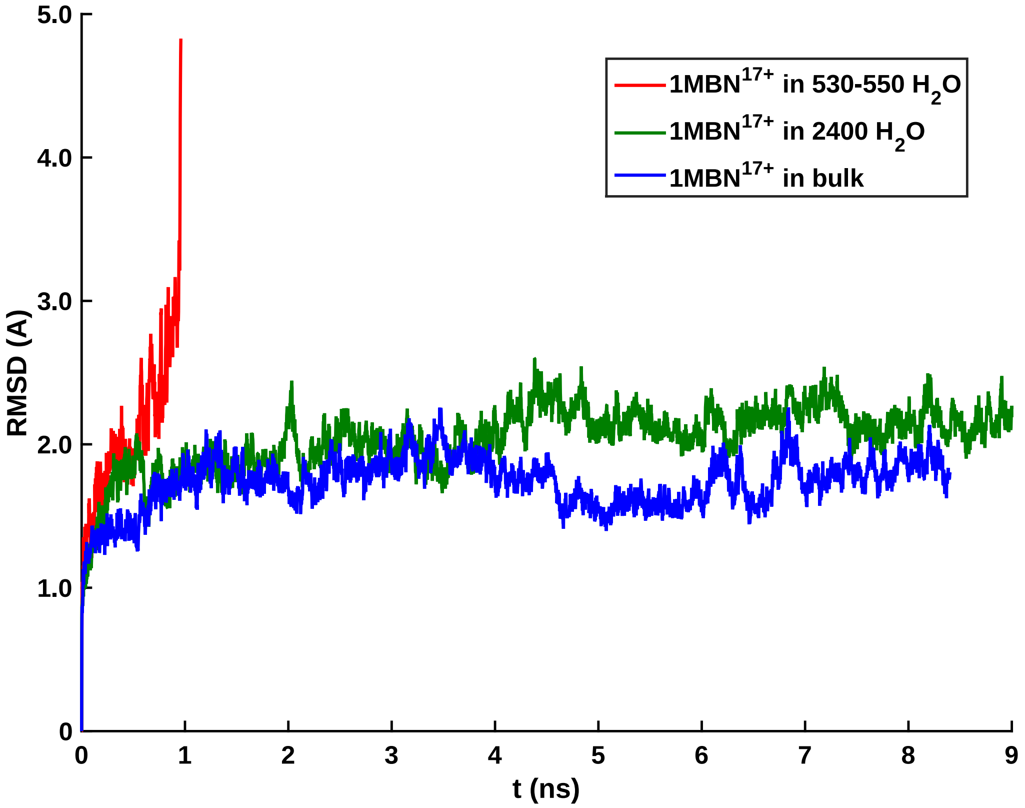
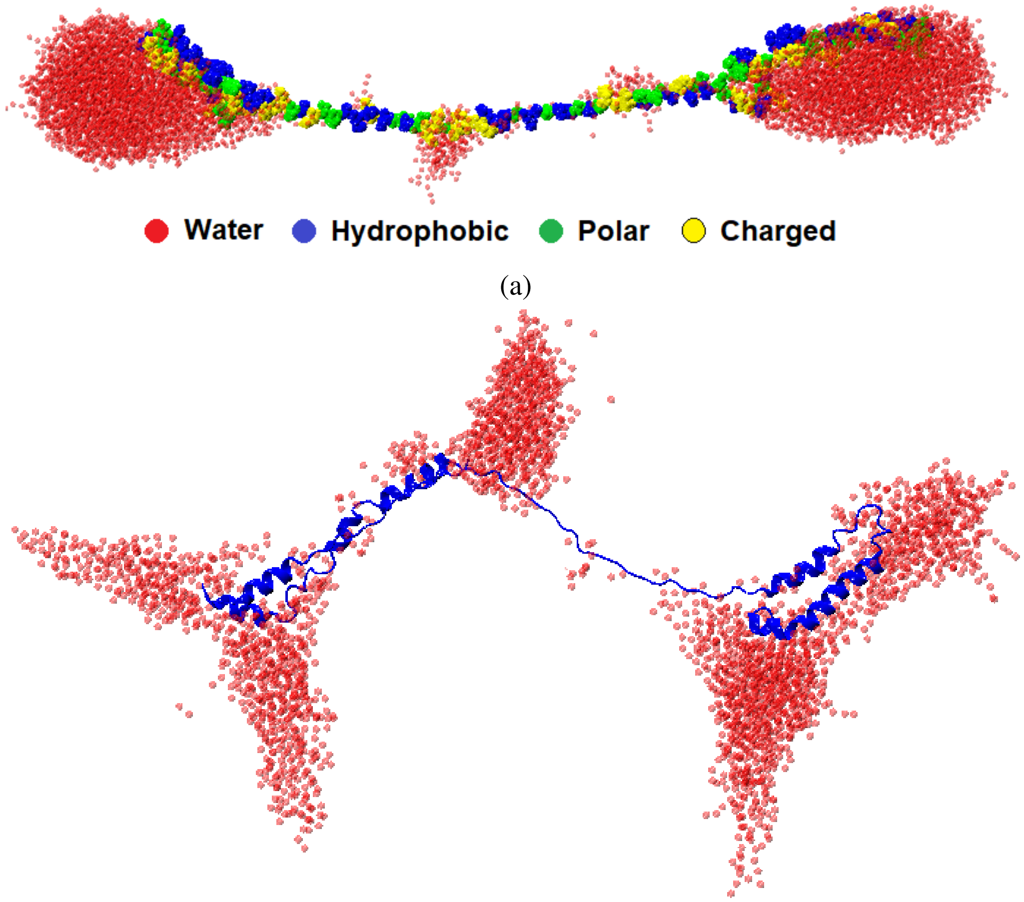


Figure 2.19: Root-mean-square displacement (RMSD) of 1MBN^{17+} in a water droplet above the Rayleigh limit (red), in a water droplet at the Rayleigh limit (green), and in bulk solution (blue).

the water is less than in the larger droplet, another form of instability appears where the solvent forms conical protrusions (see the discussion in Figure 2.22). This is the first time we observed a transition from the “pearl-necklace” morphology (Figure 2.22 (b)) to the “star” shape (Figure 2.22 (c)) .

The accumulation of the water in the chain termini is a general motif that appears not only in the presented simulations but also in simulations of a model charged macromolecule that has the backbone of PEG in a methanol droplet [94] and a charged double-stranded DNA in an aqueous droplet [11, 95]. We think that the lobes forming at the termini of $2LJL^{15+}$, $1MBN^{36+}$, and $1MBN^{17+}$ belong to the scenario shown in Figure 2.22 (b). In the case of the charged dsDNA and $1MBN^{36+}$ in an aqueous droplet of ≈ 2900 H_2O molecules (shown in Figure 2.21 (b)), the systems are substantially above the Rayleigh limit, and for this reason, conical solvent extrusions are formed at the termini. This is an intriguing finding because the water molecules do not concentrate close to the middle of the chain where the electric field is highest but they structure around the termini where the electric field is weakest. This pattern of an electric field is typical of a finite-length charged rod. As proposed by Sharawy and Consta [11, 95], the migration of the water to the chain ends can be explained as follows: The water molecules have more degrees of freedom to build a strong hydrogen bond network among them near the polypeptide termini. In contrast, the dipole moments of the water molecules is aligned with the strong electric field near the centre of the chain, restricting on their freedom to establish a hydrogen bond network. As a direct consequence of this unequal electric field strength around the chain, the H_2O molecules move to the termini of the polypeptide chain. Also, that the water is accumulated at the termini in order to disperse the charge of the chain at a larger distance, thus reducing the electrostatic energy of the system.

The overall picture that emerges about the manner in which a protein is desolvated is shown schematically in Figure 2.22. A denatured protein or an IDP may obtain a variety of conformations in bulk aqueous solution and in a microscopic droplet (Figure 2.22 (a)). In a nanoscopic droplet, a protein does not extrude (Figure 2.22 (b)), but it may extend because of charge-induced instability. Initially, the water forms pear-like lobes in the protein termini (Figure 2.22 (c)). As the water molecules continue to evaporate, this form of instability progresses into the formation of multi-point stars (Figure 2.22 (d)). Finally, the protein dries out (Figure 2.22 (e)).



(a)

(b)

Figure 2.20: (a) Equilibrium state of the extended protein 1MBN^{36+} in an aqueous droplet of approximately 6100 H_2O molecules. The nature of the amino acids throughout the chain is shown. (b) 1MBN^{36+} in an aqueous droplet of $\approx$ 2750 H_2O molecules. The nature of the amino acids is the same as in (a).

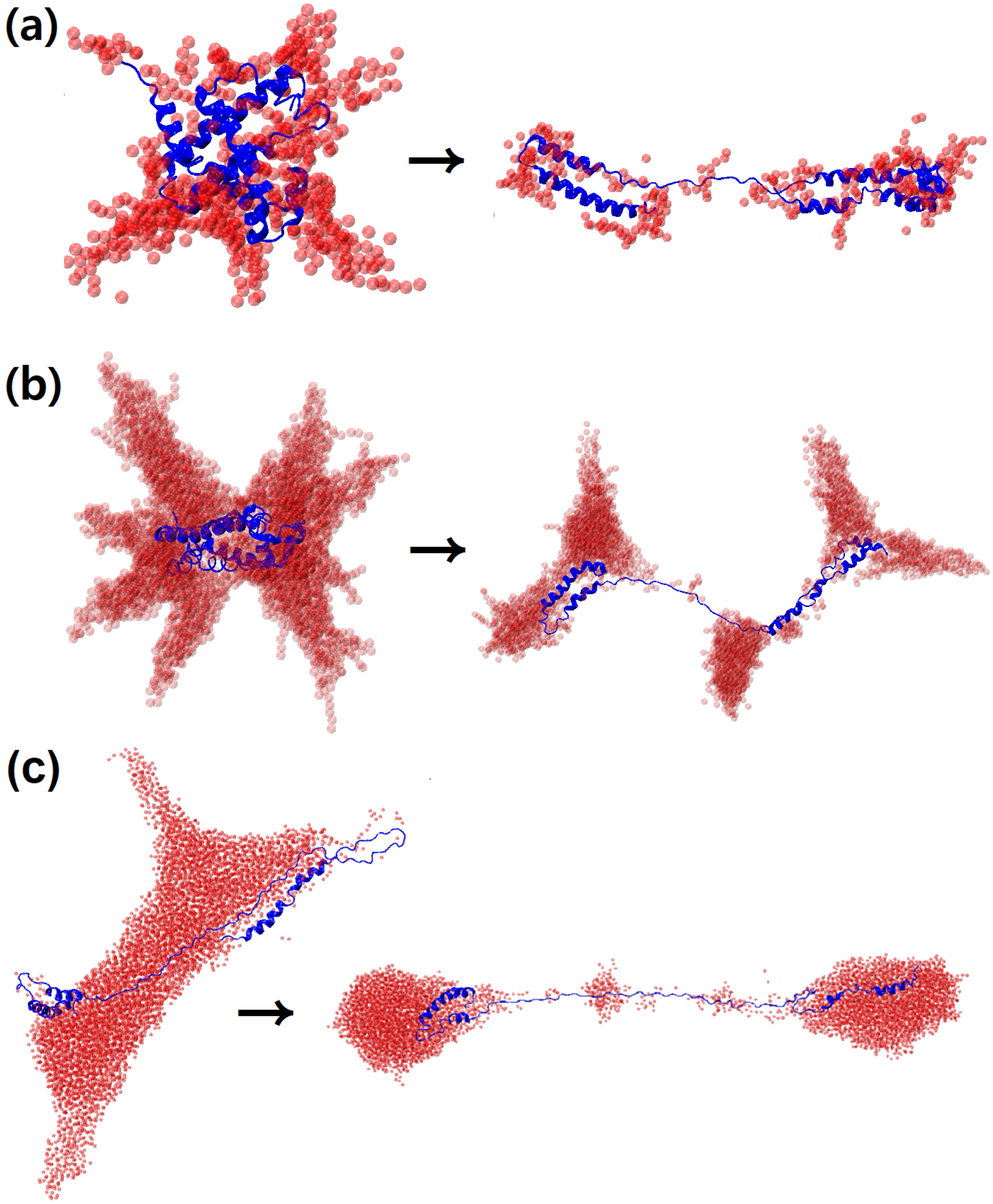


Figure 2.21: Charge-induced unfolding of (a) $1MBN^{17+}$ in an aqueous droplet of 530−550 H_2O molecules. (b) $1MBN^{36+}$ (initially compact) in an aqueous droplet of $\approx$ 2900 H_2O molecules after equilibration. (c) $1MBN^{36+}$ (initially prepared unfolded) in an aqueous droplet of $\approx$ 6177 H_2O molecules after equilibration. The same colour scheme is the same as in Figure 2.17.

The comparison between the examples of protein-aqueous droplet interactions and that of the extrusion of PEG suggests that it is very unlikely a protein to be extruded from an aqueous droplet of radius of several nanometers. One of the factors that supports this statement is the persistently encountered motif (with different solvents and macroions) of the solvent forming droplets at the termini of a macroion. In an experimental set-up where the proteins are sprayed from a physiological solution, the protonation of the amino acids will most likely occur gradually inside the droplet. The solvation of the gradually protonated protein will be different from that of a fully protonated protein that may be produced in a highly acidic solution. However, based on the consistent findings, the increased solvation of a fully protonated protein will prevent even more than a possible extrusion mechanism.

2.3.3 Comparisons with Experimental Observations

Now it is interesting to compare results from this study with the charge ejection model (CEM). The CEM has been proposed to associate the conformations of unfolded proteins with their high charge states in an electrospray ionization (ESI) process [25]. In this model, it has been claimed that a combination of electrostatic repulsion and hydrophobicity drive unfolded proteins to appear in an extended conformation from the surface of an aqueous droplet, in analogy to collision-induced dissociation of a gaseous multi-protein complex. The high charge state of an extended protein is attributed to proton equilibration between the electrosprayed droplet and the protein that is being expelled.

Konermann and his co-workers [25] state that the CEM applies to macromolecules that are (1) disordered in structure, (2) partially hydrophobic, and (3) capable of bearing excess charges. In order to verify the CEM, the authors performed molecular simulations of a spring-bead model in a nanoscopic droplet and compared with experimental results. However, our simulations show explicitly that a highly charged protein in a nanoscopic droplet under the stringent charge conditions above the Rayleigh limit does not follow chain extrusion, regardless of its conformations. Instead, the system is stabilized by splitting the droplet into two water lobes located at the ends of the macromolecule. The model that was used for protein modelling by Konermann and coworkers was a spring-bead model that is not parametrized to account for

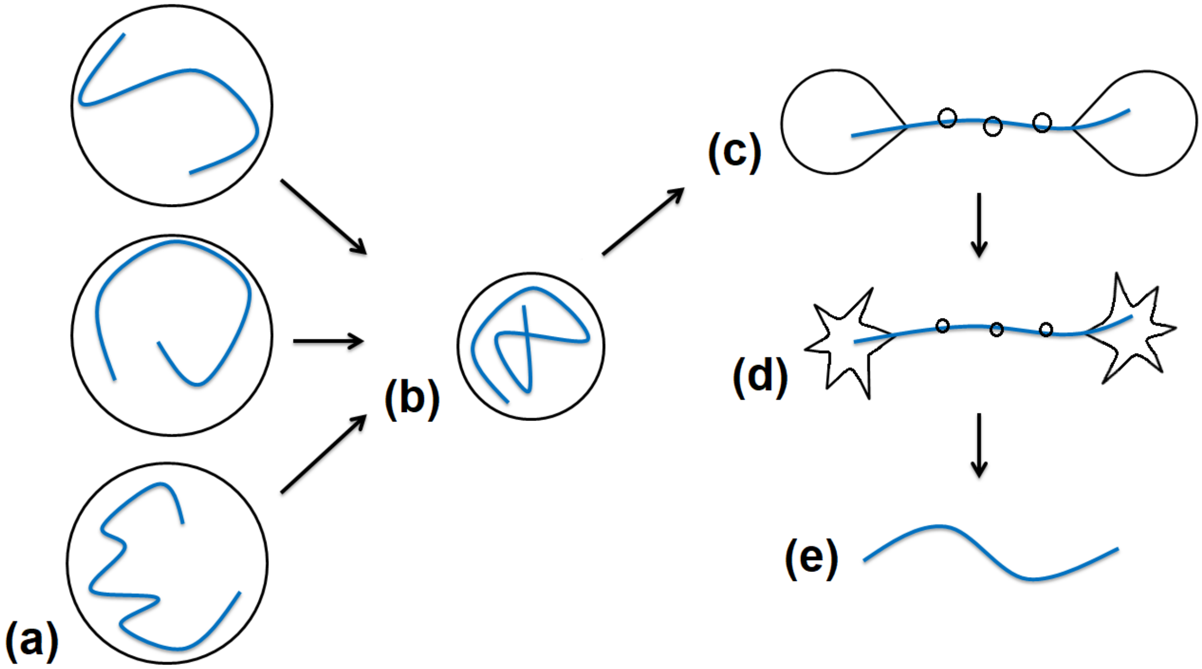


Figure 2.22: Schematic picture of the stages that are followed in the release of a charged unstructured protein from an aqueous droplet. (a) An unstructured protein (denatured or IDP) in a large droplet. The protein may have different conformations depending on its degree of disorder [32]. (b) In a nanoscopic droplet the protein may be still compact or coiled solvated by the droplet. Hydrophobic segments can be found on the surface of the droplet. (c) If the protein is highly charged, then charge-induced instability causes the stretching of the protein and the division of the water into two sub-droplets surrounding the termini of the chain. This is a stable droplet morphology under certain conditions of temperature and vapour pressure. Each sub-droplet has charge below the Rayleigh limit. Along the middle part of the chain, there may be still small aqueous islands solvating polar and charged amino acids. (d) Further solvent evaporation leads to the formation of two "star"-shaped droplets located at the termini of the chain. (e) The chain dries out.

the water-protein interactions, therefore this model could not resolve the contact points of a protein with water as we discussed in previous sections. For verification of the CEM, Ahadi and Konermann [26] performed ESI-MS experiments with horse heart myoglobin in aqueous solution at pH 7 in the presence of ammonium acetate and in aqueous solution acidified to pH 2 with formic acid. The authors ascribed enhanced signal intensities in ESI mass spectra of unfolded myoglobin to its higher desolvation and ejection rates, and claimed that these are in qualitative agreement with the CEM mechanism of an unfolded hydrophobic spring-bead chain model. We think that these observations do not provide direct evidence of the CEM. For example, the MS results may be also explained by enhanced solvent evaporation due to the increase in droplet surface area arising from the star formation of a highly charged droplet containing the macroion [11, 95–97] and the stretching of the macroion followed by the division of the droplet into two lobes as we showed.

2.4 Conclusions

In this chapter, we identified two mechanisms of charge-induced droplet disintegration in the presence of a flexible linear macromolecule, aiming to better understand how the macroion-droplet interactions determine the droplet shape and thus the final charge state of the macroion. A direct atomistic MD method provides molecular-level descriptions that macroion-droplet interactions play a critical role in explaining the manners that the macorion appears from a charged droplet as well as the origin and the magnitude of the charge state of a detected macroion in ESI-MS experiments.

The first section of the chapter focuses on how PEGs are sodiated and released from charged droplets of different solvents. The key findings are: (1) In water, PEG is localized onto the droplet surface where its segments capture Na^+ ions and extrude into the vapour phase via the CI-EM. In MeCN, in contrast, PEG resides inside the droplet, eventually emerging in the vapour phase (with bound Na^+ ions) after nearly all MeCN has evaporated away. (2) The different solvent-PEG interactions determine the final charge state and thus the conformation of PEG. Almost always PEG64-$4Na^+$ is found to be extruded from an aqueous droplet, whereas

PEG64 bound to fewer than four Na^+ ions is the most probable outcome from an MeCN droplet. Depending on the final charge state, the released macromolecule is compact, partially stretched, and extensively stretched [17]. Interestingly, at the latest stage of droplet desolvation, a few MeCN molecules suffice to maintain the charged macromolecule compact. (3) Although PEG displays distinct behaviour of solvation and sodiation in droplets of pristine H_2O and MeCN solvents, when PEG is in a droplet composed of a binary mixture of the two solvents, the final charge state is almost always four as in a pristine water droplet. This is attributed to (a) the preferential solvation of Na^+ ions in water and (b) the aqueous core formation due to differential solvent evaporation of water and MeCN. Therefore, the nature of solvent can grant complexity to solvent-ion-macromolecule interactions, which, in turn, determines the final charge state and conformation of a macromolecule. Moreover, based on the simulation data, an analytical model can be constructed, which suggests that the droplet surface electric field may play a role in the macroion-droplet interactions that lead to the extrusion of the macroion. Although the role of the surface electric field has been suggested previously [98, 99], this study provides the first evidence of the effect of the surface electric field via molecular simulations with atomistic details.

In the second section of this chapter, we investigated the charged protein-aqueous droplet interactions of examples of proteins from the group of membrane proteins, IDPs, and denatured proteins. A membrane protein is composed mainly of hydrophobic amino acids and it has low charge. As expected, this protein resides on the surface of an aqueous droplet. By comparing the solvation of the membrane protein to that of an atomistically modelled PEG in an aqueous droplet, it is revealed that a membrane protein has many more hydrogen bonds with water than PEG. Even though a membrane protein is hydrophobic and carries some charge in its flexible tail, it is rather unlikely to be extruded from a highly charged droplet. A highly charged IDP and a highly charged denatured protein demonstrate common features in their interactions with water molecules in a droplet. The water molecules accumulate at the termini of the chains forming two lobes. Even though it is observed in non-equilibrium states that the water may surround a protein almost uniformly, this is only transient and the water rapidly shrinks to form the two lobes. It appears as in this process the water is expelled by the body of the macroion although polar and charged amino acids are everywhere distributed throughout the macroion. The stable droplet-protein morphology is consistent with our “pearl-necklace” scenario. Accumulated evidence from PEG in methanol and the present simulations shows that the lobes are not spherical but they have a “pear” shape. We think that the water is accumulated at the termini so as to separate the charge of the chain at a larger distance and so reduce the electrostatic energy. The central region of the chain is exposed to the vapour, yet this is not because of chain extrusion but because of the “pearl-necklace” form of instability. Therefore, the extrusion mechanism of a chain is very different from the unfolding of a chain due to its high charge. It is also found that when the droplet deviates even more from the Rayleigh limit in the instability regime, the instability changes its form by developing conical protrusions of the solvent. This is the first evidence that shows transitions between forms of instability in the same system.

We expect that linear macromolecules with the solvation properties of PEG will also be released via an extrusion mechanism. The study of a possible extrusion mechanism of a protein from a droplet is a very challenging question. Differently from PEG, which is a flexible macroion with well defined conformation on the droplet surface or interior, proteins may have a certain structure in an aqueous solution even when they are found in highly charge states. Thus, there are two obstacles to tackle in the study: (a) the force fields have not been optimized to account for the various protein conformations in the various charge states besides the native charge state and (b) the sampling of the protein conformations in solution (droplet or bulk) is very demanding at best. The droplet environment has different acidity from that of the parent bulk solution, and therefore, the protonation state of the protein may change within the droplet [8], depending on the dynamics of the various processes (proton transfer reactions, solvent and ion evaporation). Whether the protein conformation will change within the droplet lifetime is another challenging question. To address this question one possible way is to implement the multiscale methodology that we devised to study the stability of protein complexes in Chapter 4. In a simplified model that does not consider the deprotonation of a protein, the conformation of a protein may change dramatically in a small droplet due to a charge-induced instability that may cause the extension of the chain. This process is similar to the dissociation of the DNA [95], or that of the fission of a droplet due to a high charge. We think that a very likely solvent distribution along the backbone of a charged protein is that where each termini of the extended protein is surrounded by a droplet. This conformation also appears in a highly charged PEG in methanol, therefore it is a general solvation pattern [94] that may appear in general in the solvation of a charged finite rod of certain length. An extended conformation of a protein like that of 2LJL^{15+} may result from a more compact protein that may undergo a charge-induced unfolding in a droplet.

References

[1] Felitsyn, N., Peschke, M. & Kebarle, P. Origin and number of charges observed on multiply-protonated native proteins produced by ESI. *Int. J. Mass Spectrom.* **219**, 39–62 (2002).

[2] Barylyuk, K., Gülbakan, B., Xie, X. & Zenobi, R. DNA oligonucleotides: A model system with tunable binding strength to study monomer-dimer equilibria with electrospray ionization-mass spectrometry. *Anal. Chem.* **85**, 11902–11912 (2013).

[3] Servage, K. A., Silveira, J. A., Fort, K. L., Clemmer, D. E. & Russell, D. H. Water-mediated dimerization of ubiquitin ions captured by cryogenic ion mobility-mass spectrometry. *J. Phys. Chem. Lett.* **6**, 4947–4951 (2015).

[4] de la Mora, J. F. Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Anal. Chim. Acta* **406**, 93–104 (2000).

[5] Tolić, L. P. *et al.* Electrospray ionization fourier transform ion cyclotron resonance mass spectrometric characterization of high molecular mass starburst (TM) dendrimers. *Int. J. Mass Spectrom. Ion Process.* **165-166**, 405–418 (1997).

[6] Heck, A. J. R. & van den Heuvel, R. H. H. Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom. Rev.* **23**, 368–389 (2004).

[7] Catalina, M. I., van den Heuvel, R. H. H., van Duijn, E. & Heck, A. J. R. Decharging of globular proteins and protein complexes in electrospray. *Chem. Eur. J.* **11**, 960–968 (2005).

[8] Malevanets, A. & Consta, S. Variation of droplet acidity during evaporation. *J. Chem. Phys.* **138**, 184312 (2013).

[9] Nohmi, T. & Fenn, J. Electrospray mass spectrometry of poly(ethylene glycols) with molecular weights up to five million. *J. Am. Chem. Soc.* **114**, 3241–3246 (1992).

[10] Fenn, J. B., Rosell, J., Nohmi, T., Shen, S. & Banks, F. J. *Electrospray Ion Formation: Desorption Versus Desertion*, chap. 3, 60–80 (American Chemical Society, 1996).

[11] Sharawy, M. & Consta, S. Characterization of "star" droplet morphologies induced by charged macromolecules. *J. Phys. Chem. A* **120**, 8871–8880 (2016).

[12] Oh, M. I., Malevanets, A., Paliy, M., Frenkel, D. & Consta, S. When droplets become stars: charged dielectric droplets beyond the Rayleigh limit. *Soft Matter* **13**, 8781–8795 (2017).

[13] Nguyen, H. P. & Schug, K. A. The advantages of ESI-MS detection in conjunction with hilic mode separations: Fundamentals and applications. *J. Sep. Sci.* **31**, 1465–1480 (2008).

[14] Iavarone, A. T. & Williams, E. R. Mechanism of charging and supercharging molecules in electrospray ionization. *J. Am. Chem. Soc.* **125**, 2319–2327 (2003).

[15] Kienberger, F. *et al.* Static and dynamical properties of single poly(ethylene glycol) molecules investigated by force spectroscopy. *Single Mol.* **1**, 123–128 (2000).

[16] Lee, H., Venable, R. M., MacKerell, A. D. & Pastor, R. W. Molecular dynamics studies of polyethylene oxide and polyethylene glycol: Hydrodynamic radius and shape anisotropy. *Biophys. J.* **95**, 1590–1599 (2008).

[17] Consta, S. & Chung, J. K. Charge-induced conformational changes of PEG-($Na_n{}^+$) in a vacuum and aqueous nanodroplets. *J. Phys. Chem. B* **115**, 10447–10455 (2011).

[18] Chung, J. K. & Consta, S. Release mechanisms of poly(ethylene glycol) macroions from aqueous charged nanodroplets. *J. Phys. Chem. B* **116**, 5777–5785 (2012).

[19] Wong, S. F., Meng, C. K. & Fenn, J. B. Multlple charging in electrospray ionization of poly(ethylene glycols). *J. Chem. Phys.* **92**, 546–550 (1988).

[20] Ude, S., de la Mora, J. F. & Thomson, B. A. Charge-induced unfolding of multiply charged polyethylene glycol ions. *J. Am. Chem. Soc.* **126**, 12184–12190 (2004).

[21] Larriba, C. & de la Mora, J. F. The gas phase structure of coulombically stretched polyethylene glycol ions. *J. Phys. Chem. B* **116**, 593–598 (2012).

[22] Bogan, M. J. & Agnes, G. R. Poly(ethylene glycol) doubly and singly cationized by different alkali metal ions: Relative cation affinities and cation-dependent resolution in a quadrupole ion trap mass spectrometer. *J. Am. Soc. Mass Spectrom.* **13**, 177–186 (2002).

[23] von Helden, G., Wyttenbach, T. & Bowers, M. T. Conformation of macromolecules in the gas phase: use of matrix-assisted laser desorption methods in ion chromatography. *Science* **267**, 1483 (1995).

[24] Gidden, J., Wyttenbach, T., Jackson, A. T., Scrivens, J. H. & Bowers, M. T. Gas-phase conformations of synthetic polymers: poly(ethylene glycol), poly(propylene glycol), and poly(tetramethylene glycol). *J. Am. Chem. Soc.* **122**, 4692–4699 (2000).

[25] Konermann, L., Ahadi, E., Rodriguez, A. D. & Vahidi, S. Unraveling the mechanism of electrospray ionization. *Anal. Chem.* **85**, 2–9 (2013).

[26] Ahadi, E. & Konermann, L. Modeling the behavior of coarse-grained polymer chains in charged water droplets: Implications for the mechanism of electrospray ionization. *J. Phys. Chem. B* **116**, 104–112 (2012).

[27] Mehmood, S., Allison, T. M. & Robinson, C. V. Mass spectrometry of protein complexes: From origins to applications. *Annu. Rev. Phys. Chem.* **66**, 453–474 (2015).

[28] Consta, S. Manifestation of Rayleigh instability in droplets containing multiply charged macroions. *J. Phys. Chem. B* **114**, 5263–5268 (2010).

[29] Uversky, V. N. Intrinsically disordered proteins from A to Z. *Int. J. Biochem. Cell Biol.* **43**, 1090–1103 (2011).

[30] van der Lee, R. *et al.* Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **114**, 6589–6631 (2014).

[31] Oldfield, C. J. & Dunker, A. K. Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu. Rev. Biochem.* **83**, 553–584 (2014).

[32] Das, R. K. & Pappu, R. V. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 13392–13397 (2013).

[33] Stirnemann, G., Giganti, D., Fernandez, J. M. & Berne, B. J. Elasticity, structure, and relaxation of extended proteins under force. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 3847–3852 (2013).

[34] Chandrasekhar, J., Spellmeyer, D. C. & Jorgensen, W. L. Energy component analysis for dilute aqueous-solutions of Li^+, Na^+, F^-, and Cl^- ions. *J. Am. Chem. Soc.* **106**, 903–910 (1984).

[35] Jorgensen, W. L. & Tirado-Rives, J. The opls force field for proteins. energy minimizations for crystals of cyclic peptides and Crambin. *J. Am. Chem. Soc.* **110**, 1657–1666 (1988).

[36] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

[37] Lindahl, E., Hess, B. & van der Spoel, D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Mod.* **7**, 306–317 (2001).

[38] Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **14**, 33–38 (1996).

[39] O'Neil, M. J. *The Merck Index* (The Royal Society of Chemistry, 2013).

[40] Sharawy, M. & Consta, S. Effect of counterions on the charging mechanisms of a macromolecule in aqueous nanodrops. *J. Chem. Phys.* **141**, 104321 (2014).

[41] Cerdà, J. J., Sintes, T. & Chakrabarti, A. Excluded volume effects on polymer chains confined to spherical surfaces. *Macromolecules* **38**, 1469–1477 (2005).

[42] Hezaveh, S., Samanta, S., Milano, G. & Roccatano, D. Structure and dynamics of 1,2-dimethoxyethane and 1,2-dimethoxypropane in aqueous and non-aqueous solutions: A molecular dynamics study. *J. Chem. Phys.* **135**, 164501 (2011).

[43] Hezaveh, S., Samanta, S., Milano, G. & Roccatano, D. Molecular dynamics simulation study of solvent effects on conformation and dynamics of polyethylene oxide and polypropylene oxide chains in water and in common organic solvents. *J. Chem. Phys.* **136**, 124901 (2012).

[44] Kim, M. W. Surface activity and property of polyethyleneoxide (PEO) in water. *Colloids Surf. A* **128**, 145–154 (1997).

[45] Israelachvili, J. The different faces of poly(ethylene glycol). *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8378–8379 (1997).

[46] Prasitnok, K. & Wilson, M. R. A coarse-grained model for polyethylene glycol in bulk water and at a water/air interface. *Phys. Chem. Chem. Phys.* **15**, 17093–17104 (2013).

[47] Vega, C. & de Miguel, E. Surface tension of the most popular models of water by using the test-area simulation method. *J. Chem. Phys.* **126**, 154707 (2007).

[48] Chen, F. & Smith, P. E. Simulated surface tensions of common water models. *J. Chem. Phys.* **126**, 221101 (2007).

[49] Sun, L. *et al.* Molecular dynamics simulations of the surface tension and structure of salt solutions and clusters. *J. Phys. Chem. B* **116**, 3198–3204 (2012).

[50] Consta, S. & Malevanets, A. Manifestations of charge induced instability in droplets effected by charged macromolecules. *Phys. Rev. Lett.* **109**, 148301 (2012).

[51] Caleman, C., Hub, J. S., van Maaren, P. J. & van der Spoel, D. Atomistic simulation of ion solvation in water explains surface preference of halides. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 6838–6842 (2011).

[52] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (2013). URL `http://www.R-project.org/`.

[53] Nasibulin, A. G., de la Mora, J. F. & Kauppinent, E. I. Ion-induced nucleation of dibutyl phthalate vapors on spherical and nonspherical singly and multiply charged polyethylene glycol ions. *J. Phys. Chem. A* **112**, 1133–1138 (2008).

[54] Korosi, G. & Kovats, E. S. Density and surface tension of 83 organic liquids. *J. Chem. Eng. Data* **26**, 323–332 (1981).

[55] Criado-Hidalgo, E., Fernández-García, J. & Fernández de la Mora, J. Mass and charge distribution analysis in negative electrosprays of large polyethylene glycol chains by ion mobility mass spectrometry. *Anal. Chem.* **85**, 2710–2716 (2013).

[56] Marcus, Y. The structure of and interactions in binary acetonitrile + water mixtures. *J. Phys. Org. Chem.* **25**, 1072–1085 (2012).

[57] Mountain, R. D. Microstructure and hydrogen bonding in water-acetonitrile mixtures. *J. Phys. Chem. B* **114**, 16460–16464 (2010).

[58] Oldiges, C., Wittler, K., Tonsing, T. & Alijah, A. MD calculated structural properties of clusters in liquid acetonitrile/water mixtures with various contents of acetonitrile. *J. Phys. Chem. A* **106**, 7147–7154 (2002).

[59] Shin, D. N., Wijnen, J. W., Engberts, J. B. & Wakisaka, A. On the origin of microheterogeneity: Mass spectrometric studies of acetonitrile-water and dimethyl sulfoxide-water binary mixtures (part 2). *J. Phys. Chem. B* **106**, 6014–6020 (2002).

[60] Takamuku, T. *et al.* Liquid structure of acetonitrile-water mixtures by X-ray diffraction and infrared spectroscopy. *J. Phys. Chem. B* **102**, 8880–8888 (1998).

[61] Shin, D. N., Wijnen, J. W., Engberts, J. B. & Wakisaka, A. On the origin of microheterogeneity: A mass spectrometric study of dimethyl sulfoxide-water binary mixture. *J. Phys. Chem. B* **105**, 6759–6762 (2001).

[62] Cringus, D., Yeremenko, S., Pshenichnikov, M. S. & Wiersma, D. A. Hydrogen bonding and vibrational energy relaxation in water-acetonitrile mixtures. *J. Phys. Chem. B* **108**, 10376–10387 (2004).

[63] Marcus, Y. Preferential solvation of ions in mixed solvents. Part 2. - the solvent composition near the ion. *J. Chem. Soc., Faraday Trans. 1* **84**, 1465–1473 (1988).

[64] Stellnberger, K.-H., Gritzner, G., Lewandowski, A. & Orlik, M. Thermodynamic transfer properties for Na^+, Tl^+ and Ag^+ to water-acetonitrile mixtures. *J. Chem. Soc., Faraday Trans.* **91**, 875–880 (1995).

[65] Covington, A. K. & Dunn, M. Nuclear magnetic resonance studies of preferential solvation. Part 7. - sodium iodide in ethylene glycol-acetonitrile and in propylene glycol-acetonitrile mixtures. *J. Chem. Soc., Faraday Trans. 1* **85**, 2835–2846 (1989).

[66] Ansell, S., Barnes, A., Mason, P., Neilson, G. & Ramos, S. X-ray and neutron scattering studies of the hydration structure of alkali ions in concentrated aqueous solutions. *Biophys. Chem.* **124**, 171–179 (2006).

[67] Rempe, S. B. & Pratt, L. R. The hydration number of Na^+ in liquid water. *Fluid Phase Equilib.* **183-184**, 121–132 (2001).

[68] Varma, S. & Rempe, S. B. Coordination numbers of alkali metal ions in aqueous solutions. *Biophys. Chem.* **124**, 192–199 (2006).

[69] Nguyen, T.-N. V., Hughes, S. R. & Peslherbe, G. H. Microsolvation of the sodium and iodide ions and their ion pair in acetonitrile clusters: A theoretical study. *J. Phys. Chem. B* **112**, 621–635 (2008).

[70] Alberti, M., Amat, A., De Angelis, F. & Pirani, F. A model potential for acetonitrile: from small clusters to liquid. *J. Phys. Chem. B* **117**, 7065–7076 (2013).

[71] Vaden, T. D. & Lisy, J. M. Competing non-covalent interactions in alkali metal ion-acetonitrile-water clusters. *J. Phys. Chem. A* **109**, 3880–3886 (2005).

[72] Sterling, H. J. & Williams, E. R. Origin of supercharging in electrospray ionization of noncovalent complexes from aqueous solution. *J. Am. Soc. Mass Spectrom.* **20**, 1933–1943 (2009).

[73] Šamalikova, M. & Grandori, R. Protein charge-state distributions in electrospray-ionization mass spectrometry do not appear to be limited by the surface tension of the solvent. *J. Am. Chem. Soc.* **125**, 13352–13353 (2003).

[74] Brahim, B., Alves, S., Cole, R. B. & Tabet, J.-C. Charge enhancement of single-stranded dna in negative electrospray ionization using the supercharging reagent meta-nitrobenzyl alcohol. *J. Am. Soc. Mass Spectrom.* **24**, 1988–1996 (2013).

[75] Šamalikova, M. & Grandori, R. Testing the role of solvent surface tension in protein ionization by electrospray. *J. Mass Spectrom.* **40**, 503–510 (2005).

[76] Lomeli, S. H., Yin, S., Ogorzalek Loo, R. R. & Loo, J. A. Increasing charge while preserving noncovalent protein complexes for ESI-MS. *J. Am. Soc. Mass Spectrom.* **20**, 593–596 (2009).

[77] Lomeli, S. H., Peng, I. X., Yin, S., Ogorzalek Loo, R. R. & Loo, J. A. New reagents for increasing ESI multiple charging of proteins and protein complexes. *J. Am. Soc. Mass Spectrom.* **21**, 127–131 (2010).

[78] Ogorzalek Loo, R. R., Lakshmanan, R. & Loo, J. A. What protein charging (and supercharging) reveal about the mechanism of electrospray ionization. *J. Am. Soc. Mass Spectrom.* **25**, 1675–1693 (2014).

[79] Maximino, R. B. Surface tension and density of binary mixtures of monoalcohols, water and acetonitrile: equation of correlation of the surface tension. *Phys. Chem. Liq.* **47**, 475–486 (2009).

[80] Iavarone, A. T., Jurchen, J. C. & Williams, E. R. Effects of solvent on the maximum charge state and charge state distribution of protein ions produced by electrospray ionization. *J. Am. Soc. Mass Spectrom.* **11**, 976–985 (2000).

[81] Hopper, J. T., Sokratous, K. & Oldham, N. J. Charge state and adduct reduction in electrospray ionization-mass spectrometry using solvent vapor exposure. *Anal. Biochem.* **421**, 788–790 (2012).

[82] Berman, H. M. *et al.* The protein data bank. *Nucleic Acids Res.* **28**, 235–242 (2000).

[83] `http://www.rcsb.org`.

[84] Barrera-Vilarmau, S., Obregón, P. & de Alba, E. Intrinsic order and disorder in the Bcl-2 member harakiri: Insights into its proapoptotic activity. *PLOS ONE* **6**, 1–13 (2011).

[85] Singarapu, K. K. *et al.* Structural characterization of Hsp12, the heat shock protein from saccharomyces cerevisiae, in aqueous solution where it is intrinsically disordered and in detergent micelles where it is locally α-helical. *J. Biol. Chem.* **286**, 43447–43453 (2011).

[86] Watson, H. C. The stereochemistry of the protein myoglobin. *Prog. Stereochem.* **4**, 299–333 (1969).

[87] Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comp. Chem.* **31**, 671–690 (2010).

[88] Yu, W., He, X., Vanommeslaeghe, K. & MacKerell, A. D. Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations. *J. Comp. Chem.* **33**, 2451–2468 (2012).

[89] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

[90] Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comp. Chem.* **26**, 1781–1802 (2005).

[91] Jorgensen, W. L. & Jenson, C. Temperature dependence of TIP3P, SPC, and TIP4P water from NPT monte carlo simulations: Seeking temperatures of maximum density. *J. Comp. Chem.* **19**, 1179–1186 (1998).

[92] Vega, C. & Abascal, J. L. F. Simulating water with rigid non-polarizable models: a general perspective. *Phys. Chem. Chem. Phys.* **13**, 19663–19688 (2011).

[93] Consta, S. & Malevanets, A. Disintegration mechanisms of charged nanodroplets: novel systems for applying methods of activated processes. *Mol. Simul.* **41**, 73–85 (2015).

[94] Consta, S. & Malevanets, A. Classification of the ejection mechanisms of charged macromolecules from liquid droplets. *J. Chem. Phys.* **138**, 044314 (2013).

[95] Sharawy, M. & Consta, S. How do non-covalent complexes dissociate in droplets? A case study of the desolvation of dsDNA from a charged aqueous nanodrop. *Phys. Chem. Chem. Phys.* **17**, 25550–25562 (2015).

[96] Oh, M. I. & Consta, S. Stability of a transient protein complex in a charged aqueous droplet with variable pH. *J. Phys. Chem. Lett.* **8**, 80–85 (2017).

[97] Oh, M. I. & Consta, S. What factors determine the stability of a weak protein-protein interaction in a charged aqueous droplet? *Phys. Chem. Chem. Phys.* **19**, 31965–31981 (2017).

[98] Iribarne, J. & Thomson, B. Evaporation of small ions from charged droplets. *J. Chem. Phys.* **64**, 2287–2294 (1976).

[99] Thomson, B. & Iribarne, J. Field-induced ion evaporation from liquid surfaces at atmospheric-pressure. *J. Chem. Phys.* **71**, 4451–4463 (1979).

Chapter 3

Characterization of “Star” Morphologies of Charged Droplets

3.1 Motivation and Objectives

Droplet deformations in an electric field have received a lot of attention since the dawn of the theory of electrostatics [1–7]. Starting from the seminal works of Rayleigh [1], extended by Zeleny [8] and Taylor [2], the behaviour of droplets in an electric field continues to fascinate scientists due to their numerous applications [9–11]. The electric field may be external to the droplet, or it may be generated by ions found within the droplet. This electric field (internal or external) may induce distinct droplet morphologies and structures when its strength is beyond a threshold value [2–4]. The properties of these different shapes and structures have been exploited in many applications. Examples are found in the electrohydrodynamic droplet deformations [12] (created by an external electric field) used in industrial applications such as crude oil demulsification, and in non-spherical shapes (created by ions) of electrosprayed nano- and micro-sized droplets [9–11] used in the generation of nano- and micro-particles. The latter example is representative of the far-reaching result of manipulating the electrostatic interactions within a droplet. In this example, the droplets are generated by the electrospray ionization (ESI) method and are composed of solvent, polymers, and ions. Their distinct structure and morphology result from the competing effects of polymer entanglement that tends to maintain the connectivity of the droplet and Coulomb fission that tends to divide the droplet.

Here, we study charged droplets with linear dimensions in the nanometer range by using atomistic modelling. In contrast to the studies where a neutral dielectric or conducting droplet

[5, 13] or a planar surface [2] is exposed to a uniform electric field [2, 5, 13], we focus on the case where the electric field is created by a single, highly charged macroion in the interior of the droplet. Hereafter, a charged chain or a spherical ion with size or charge beyond that of the simple metal cations or non-metal anions will be called a macroion. Importantly, since the droplet contains no other free charges, charge separation by fission is not possible. In general, the shape of the charged macroion affects the shape of the droplets. Here we focus on spherical ions.

As was demonstrated by Rayleigh [1], conducting droplets become unstable if their charge exceeds a threshold value. The details of the theory are found in Section 1.2 of the thesis and the derivation in Appendix B. When such droplets become unstable, they undergo fission. Consta [14] has found that when a droplet contains in its interior a single charged macroion (linear or spherical), the droplet may deform into a stable "thorny" shape [14], provided that the charge-squared-to-volume ratio of the spherical droplet exceeds a threshold value. The deformation of a spherical droplet into a "thorny" shape is a way for the droplet to lower its free energy.

We are interested in the "star"-shaped droplets for several reasons. Firstly, we think that they play an important role in determining the charge state of macroions in electrosprayed droplets. Examples of their role are presented in Chapter 4. Secondly, they may also have potential applications in materials science and catalysis as it is explained in the conclusion. Thirdly, the charge-induced instabilities of droplets are a fundamental question in science. Lord Rayleigh analyzed the stability of the conducting droplets. The analysis of the stability of the dielectric droplets was missing. It is interesting that without the guidance by the simulations, it was not possible for one to find the droplet deformations even with mathematical (continuum) modelling.

The objective of this chapter is to analyze in a systematic way the dependence of the star-shaped structures on the net charge of the macroion. Initially, the generalized Rayleigh criterion that extends to dielectrics is presented. We will examine some of the features of the star-shaped structures such as the number of spikes and the angles at the tips. The systems under investigation are nanoscopic droplets, each of which consists of solvent molecules and a single multiply charged spherical macroion located at the centre. Water and dimethyl sulfoxide (DMSO) were selected as examples of solvents. The magnitude and the sign of the charge of the macroion were varied.

3.1.1 Onset of Instability in Dielectric Droplets

For a point charge in the interior of a dielectric droplet one may develop, a linear continuum model that will allow for the analysis of the stability of a charged dielectric droplet. In this model, in the same way as the Rayleigh model, we consider a droplet with a fluctuating shape, that maintains its volume but it may change its surface area. In the macroscopic (continuum) modelling, the energy of a droplet is written as the sum of the electrostatic (E_{Coul}) and the surface energy (E_{surf})

$$E = E_{\text{surf}} + E_{\text{Coul}}. \tag{3.1}$$

In the analysis in the next paragraphs, we find the difference in energy between a spherical droplet and a droplet that deviates from a sphere. The surface energy is expressed as the product of the surface area (A) times the surface tension (γ). The surface is given by:

$$\rho(\theta, \phi) = R_0 + \sum_{l>0,m_l} a_{l,m_l} Y_{l,m_l}(\theta, \phi) \tag{3.2}$$

where (θ, ϕ) is the spherical angle, $\rho(\theta, \phi)$ is the distance from the centre, $Y_{l,m_l}(\theta, \phi)$ denote spherical harmonics of degree l and order m, and a_{l,m_l} is the expansion coefficient. R_0 is the $l = 0$ term in the expansion of $\rho(\theta, \phi)$ in terms of $Y_{l,m_l}(\theta, \phi)$. Following several algebraic steps [15] (Appendix B), we arrive at the following expression for the surface area

$$A = 4\pi R_0^2 + \sum_{l>0,m_l} \left|a_{l,m_l}\right|^2 + \frac{1}{2} \sum_{l>0,m_l} l(l+1)\left|a_{l,m_l}\right|^2 . \tag{3.3}$$

Following textbook electrostatics [16], the electrostatic energy (E_{diel}) of a linear (polarizable) dielectric with free charge is given by

$$E_{\text{diel}} = \frac{1}{2} \int\limits_{\mathbb{R}^3/V} d\mathbf{r}\mathbf{D} \cdot \mathbf{E} \tag{3.4}$$

where **D** is the electric displacement and **E** is the electric field caused by the free charge. (Vectors will be denoted by letters in boldface.) Equation 3.4 includes the response of the dielectric, therefore its polarization that may involve "stretching" and displacement of the dielectric molecules as we bring the free charge in the dielectric (or the dielectric into the free charge). When a spherical charge is surrounded by a dielectric with dielectric permitivity ε^I embedded into a dielectric with dielectric permitivity ε^E, Equation 3.4 yields that the energy of the system is given by:

$$E_{\text{diel}} = -\frac{1}{2} \int\limits_{\mathbb{R}^3/V} d\mathbf{r}(\varepsilon^E - \varepsilon^I)\mathbf{E} \cdot \mathbf{E}_0 \tag{3.5}$$

where $\mathbf{E}_0$ is the electric field from the charge in uniform dielectric ε^I.

The difference in energy (composed of the surface energy and the electrostatic energy as written in Equation 3.1) between a spherical shape and a deformed shape to a linear approximation is given by

$$\delta E = 2\pi R^2 \gamma \sum_{l>0,m} \left[(l-1)(l+2) - 4XS(\varepsilon^I/\varepsilon^E, l)\right] |a_{l,m}|^2. \tag{3.6}$$

In Equation 3.6, S is given by

$$S(\varepsilon, l) = \frac{(\varepsilon - 1)}{\varepsilon} \frac{l(l-1)\varepsilon - (l+1)(l+2)}{\varepsilon l + l + 1}. \tag{3.7}$$

where $\varepsilon = \varepsilon^I/\varepsilon^E$, and X is the fissility parameter given by

$$X = \frac{Q^2}{64\pi^2 \gamma \varepsilon_0 R^3}. \tag{3.8}$$

The details of the algebra that lead to Equation 3.6 are described in Ref. [17]. We find the criterion for dielectric droplet stability by equating to zero the first coefficient of $|a_{l,m_l}|^2$ in Equation 3.6 that may become unstable. The primary unstable mode that becomes unstable *vs.* the dielectric constant of the dielectric is shown in Figure 3.1. For $\varepsilon^I/\varepsilon^E > 17.2$, the $l = 2$ is the first mode to become unstable. For $7.74 < \varepsilon^I/\varepsilon^E < 17.2$, the $l = 3$ is the first mode to become unstable. When $\varepsilon^I/\varepsilon^E < 7.74$, the $l = 3$ mode is replaced by higher order harmonics as the primary unstable modes. As clearly seen in the same figure, the effect of the finite dielectric constant is very small for water, whose dielectric constant at 300 K is $\varepsilon^{l} \approx 77.3$.

3.2 Modelling and Computational Methods

Constant-energy atomistic molecular dynamics (MD) simulations were performed to investigate the star formation of a supercharged droplet of a pure solvent of water and DMSO. In order to hinder solvent evaporation and have equilibrium runs, the droplet was placed inside a cavity of radius 10-20 nm which was generated by applying the spherical boundary condition. All the MD simulations were performed by using the NAMD pacakage version 2.9 [18]. The CHARMM General Force Field (CGenFF) [19, 20] was used to model the TIP3P water, DMSO, and macroions. The Lowe-Andersen thermostat [21] was employed to thermalize the systems. The velocity Verlet algorithm was used to integrate the equations of motion, and the length of each time step was 1 fs. The surface tension, the mass density, and the dielectric constants of the TIP3P water and DMSO were obtained from Refs. [22–25] (summarized in Table 3.2). These quantities are not used in the MD simulations but in the analysis of the data. For instance, we need to know these values in order to estimate the Rayleigh limit. Electrostatic and van der Waals interactions were explicitly treated by using a switch scheme. The cutoff distance was determined such that every interaction within the main droplet was included in the force calculations.

Each system contains a spherical macroion centralized inside a droplet composed of 2148 H_2O molecules or 1831 DMSO molecules. Here an artificial ion was used as a paradigm to demonstrate the effect of inseparable excess charge without structural complexity of a macroion. This macroion can also represent different classes of realistic macroions with radial symmetry, such as globular proteins. The macroion was created by modifying the electrical charge and the size of a typical Na^+ ion with other parameters held constant. The ionic radius was increased to 5.0 Å, and the net charge varied ranging from 10 e to 45 e for H_2O and 20 e to 60 e for DMSO (where $e = 1.602 \times 10^{-19}$ C is the elementary charge). Each droplet system was equilibrated first at high temperature for > 1 ns and then at lower temperature ($T = 300$ K) for 1 ns. The high-temperature equilibration was intended to reduce the possibility of kinetic trapping of the

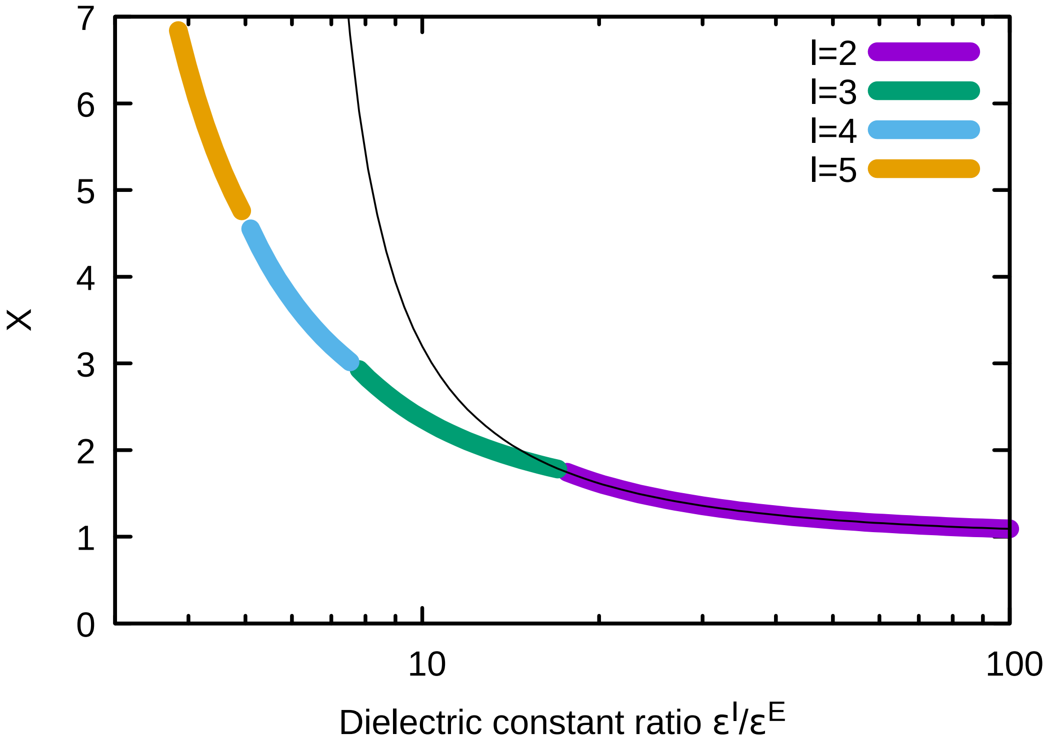


Figure 3.1: Correction to the Rayleigh criterion for media with dielectric constant ε^I. As the dielectric constant ratio decreases the primary unstable mode changes from $l = 2$ when $\varepsilon^I/\varepsilon^E > 17.2$ to $l = 3$ when $\varepsilon^I/\varepsilon^E > 7.74$ and so on. The black line was obtained from the Rayleigh criterion for comparision.

Table 3.1: Physical properties of TIP3P water and DMSO. d : mass density (in g/cm^3); γ : surface tension (in mN/m); ε_r : dielectric constant (relative permittivity); N: average number of solvent molecules in a droplet (ignoring the relatively small number of molecules that are in the vapour phase); Q_R: Rayleigh limit (in e) of the given solvent droplet; estimated Q_I critical charge (in e) corrected for the dielectric constant of the given solvent.

	TIP3P	DMSO
Formula	H_2O	$(CH_3)_2SO$
d	0.98	1.12
γ	42	49.5-53.8
ε_r	94	44
N	2148	1831
Q_R	13	22
Q_I	14	24

systems in local energy minima. After equilibration, the thermostat was switched off and the simulation continued for 6 ns or longer. To study the effect of the sign of charge on the star formation, 10 additional MD runs were also performed: −20 e, −25 e, −30 e, −35 e, and −40 e in 2148 H_2O molecules, and −20 e, −30 e, −40 e, −50 e, and −60 e in 1831 DMSO molecules.

In order to study the scalability of the number of rays with the size of the system, the MD simulations were also performed for different sizes of aqeuous droplets containing from 500 to 5000 H_2O molecules. The charge was accordingly varied from 9 e to 34 e to ensure that the fissibility parameters (Equation 1.6) are in the range of 1.6−1.8. The temperature was set to 300 K, and the lengths of the runs were 2−4 ns after 1 ns of equilibration. VMD version 1.9.2 [26] was used for visualization of the MD trajectories.

3.3 Results and Discussion

3.3.1 Stable Droplets Beyond the Instability Point

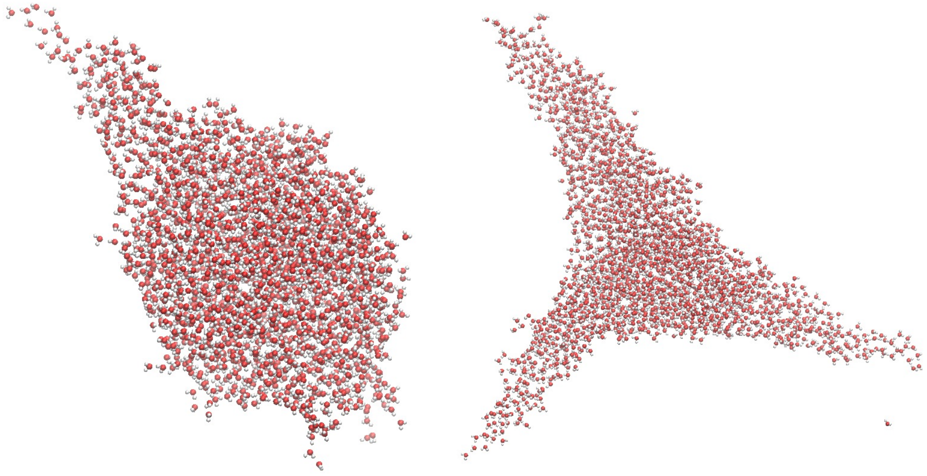
When the charge-squared-to-volume ratio of a spherical droplet exceeds a certain value predicted by Equation 3.6, a spherical droplet becomes unstable. To accommodate the instability, the unstable droplet deforms into a stable non-spherical shape that exhibits conical solvent protrusions on the droplet surface. Examples of a non-fissile macroion include a fully charged nucleic acid [27, 28] and a charged protein [14]. The conical protrusions of the solvent may appear along the backbone of a linear macroion, or they may form a three-dimensional "star" shape when the macroion is spherical. Because of the common features of the "spiky" shapes regardless of the fine details in the structure of the macroion, we will use here the example of a spherical structureless macroion sited at the centre of a droplet. A collection of typical droplet morphologies in the presence of a central macroion of various charges is shown in Figure 3.2. The charge of the corresponding spherical droplet is above the Rayleigh limit and the central charge is non-fissile. Analysis of the structure of these systems has shown that they consist of a region closest to the central macroion that is characterized by saturation of the polarization of the solvent [17, 28], and the spikes that form beyond the saturated core and show an elastic motion [28].

3.3.2 Geometry of Star-shaped Droplets

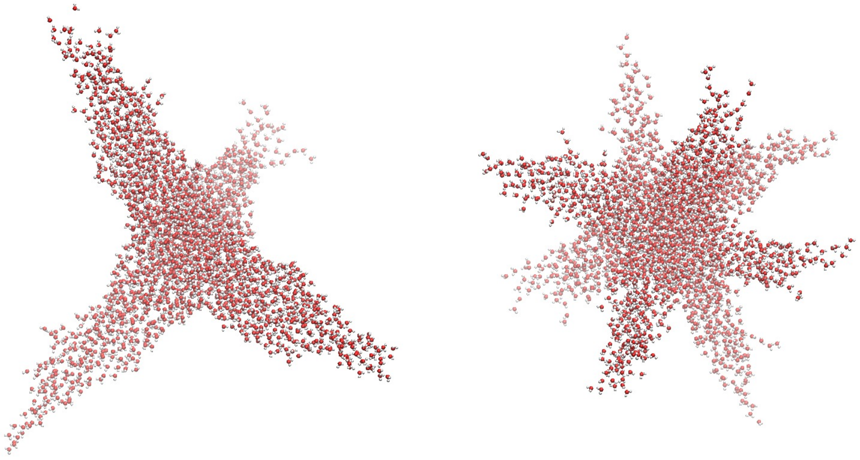
Figure 3.3 shows various stellate morphologies of H_2O (top) and DMSO (middle) droplets when their charges are above the critical values. Depending on the amount of excess charge, they have regular star shapes whose points form regular polyhedral geometries as shown at the bottom pannel.

For a droplet containing 2148 H_2O molecules, the Rayleigh limit Q_R (Equation 1.6) is estimated to be 12.9−13.4 e. This critical charge was obtained by using the surface tension value of TIP3P water (γ = 49.5−53.8 mN/m [22–24]) at T = 300 K. Likewise, for a droplet comprising 1831 DMSO molecules, the Rayleigh limit is calculated to be 21.6−22.1 e by using the surface tension value of γ = 41.5−43.3 mN/m [25] at T = 300 K. However, if we consider the dielectric behaviour of solvents in the droplets with immobile charge [17], there should be a correction factor in the Rayleigh limit as shown in Figure 3.1. The correction factors are 1.05 and 1.11 for H_2O and DMSO, respectively. These will yield Q_I =13.5−14.1 e for H_2O and Q_I =24.0−24.4 e, where Q_I denotes the critical charge corrected for dielectric droplets using Equation 3.6.

Our simulations revealed that the aqueous droplets charged with $0 < Q \leq 12$ e are in the stability regime, as menifested by their spherical shapes, whereas those with $Q \geq 14$ e become regular stellate polytopes. We found the first appearance of solvent protrusions at the surface of a DMSO droplet when Q = 25 e. Our simulations clearly show that the onset of shape



(a) 1 spike, $Q = 14\ e$ (b) 3 spikes, $Q = 18\ e$



(c) 4 spikes, $Q = 22\ e$ (d) 8 spikes, $Q = 32\ e$

Figure 3.2: Typical snapshots of charged aqueous droplets at $T = 300$ K. All the droplets are composed of 2148 H_2O molecules and have an embedded model spherical ion. The charge at the instability limit for these droplets is $Q \approx 13.9\ e$. All the droplets are charged above the Rayleigh limit.

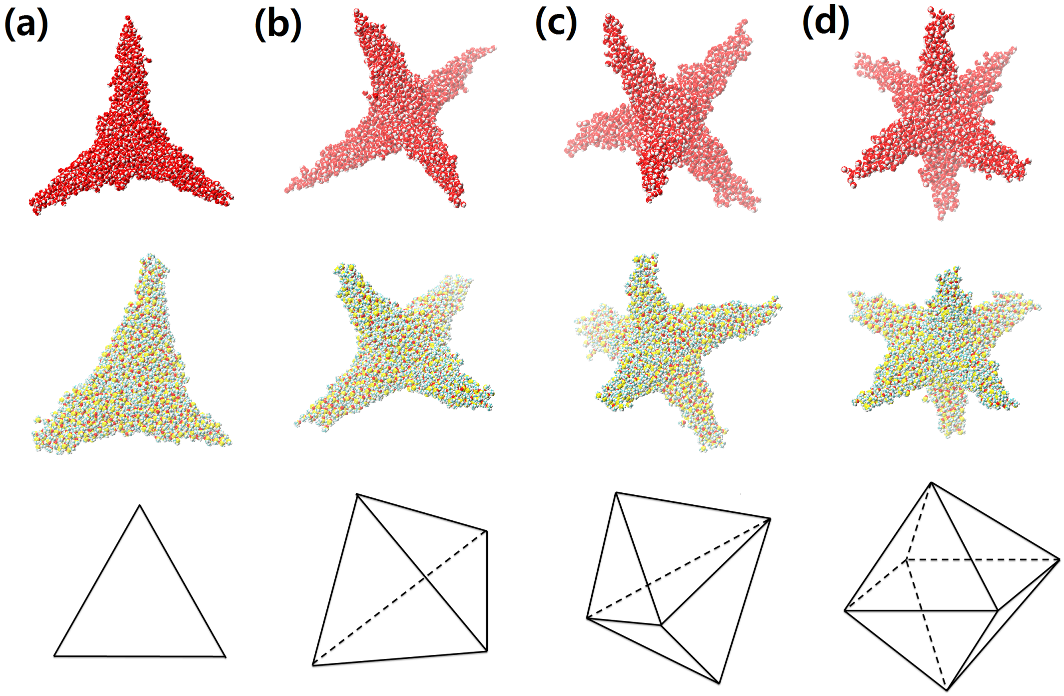


Figure 3.3: Typical snapshots of different star morphologies of droplets of 2148 H_2O (top) and 1831 DMSO (middle) molecules. The spikes form (a) a trigonal plane when $Q = 18\ e$ and $Q = 27\ e$, (b) a tetrahedron when $Q = 22\ e$ and $Q = 32\ e$, (c) a trigonal bipyramid when $Q = 25\ e$ and $Q = 35\ e$, and (d) an octahedron when $Q = 28\ e$ and $Q = 40\ e$ for H_2O and DMSO, respectively.

instability, as characterized by enhanced shape fluctuations without mature growths of spikes (Figure 3.4), occurs when $Q \approx Q_I$ (that is, $Q = 13\ e$ for H_2O and $Q = 23\ e$ for DMSO). These are in close agreement with the theory presented in Section 3.1.1 (Equation 3.6).

In the charge-induced instability regime (*i.e.*, $Q > Q_I$), the water droplets undergo shape deformations to star morphologies whose vertices attain certain polyhedral geometries of high symmetry and regularity. For instance, when $Q = 14{-}17\ e$, two arms are formed at the antipodal points of the aqueous droplet, and therefore, the conformation becomes linearly symmetric. A further increase in the net charge of the macroion leads to the birth of additional spikes on the droplet surface. The droplet is evolved into an equilateral triangle about a great circle (three spikes) when $Q = 18{-}19\ e$, a tetrahedron (four spikes) when $Q = 20{-}23\ e$, a triangular bipyramid (five spikes) when $Q = 24{-}25\ e$, an octahedron (six spikes) when $Q = 26{-}28\ e$, and an icosahedron (12 spikes) when $Q = 41{-}42\ e$. We also found that the same star shapes appear by altering the sign of the excess charge. The same behaviours were observed for the charged DMSO droplets, except that (1) they show substantial favour to the tetrahedral shape (four spikes) and (2) they require higher charge to attain the same geometry as water droplets due to their larger volumes.

It is an impressive finding that the number of angular points found in the star-shaped droplets increases with the net charge of the droplets in a quadratic manner, as shown in Figure 3.5 (dashed lines). In this graph, the number of rays N_{rays} is plotted as the fissility parameter X (Equation 3.8). This trend may be justified by geometric considerations of the droplets.

Here we present an analytical model that gives an estimate of the number of rays as a function of the charge. We denote the number of rays by N and the total amount of charge by Q. Based on the droplet morphology that we found in the simulations, we approximate the star-shaped droplet by a spherical core with a number of cones covering all the surface. Next, we approximate the charge in all the cones with a uniform distribution on a spherical shell enclosing the central core. Thus, in the model, the system is represented by two co-centric spherical shells: the interior has radius R_1 and carries charge Q_1 and the exterior has radius R_2 and charge Q_2. The distance between the shells is directly proportional to the cone height H

$$R_2 = R_1 + \beta H, \tag{3.9}$$

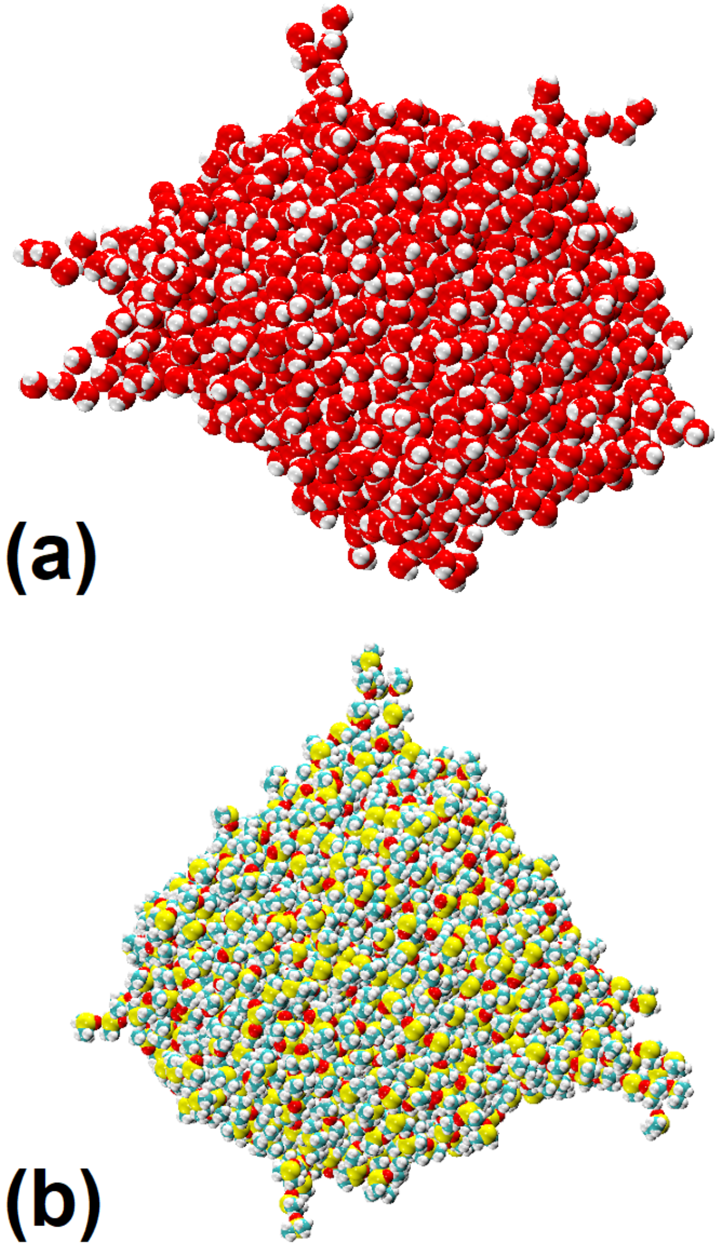


Figure 3.4: Enhanced shape fluctuations of (a) water (Q = 13 e) and (b) DMSO droplets (Q = 23 e).

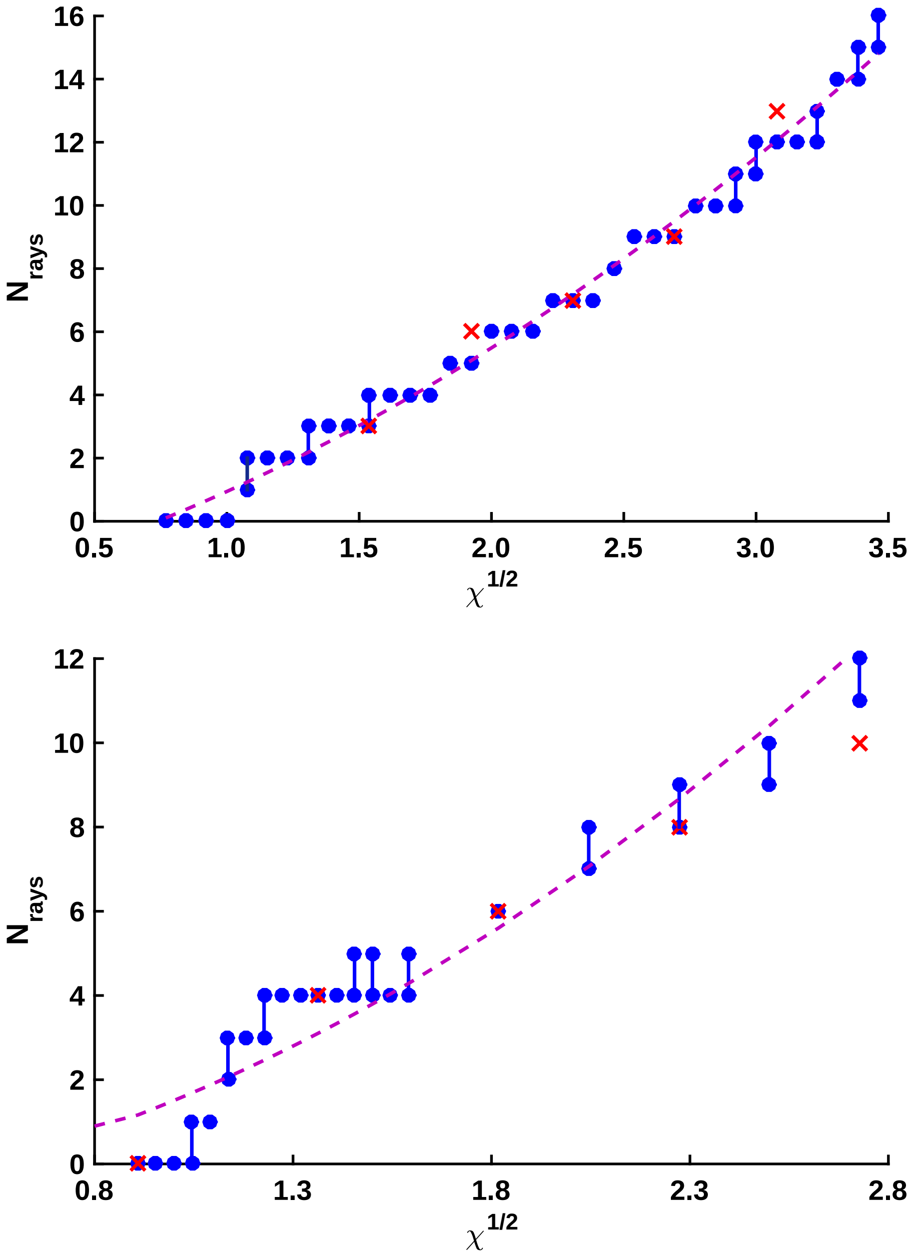


Figure 3.5: Number of points in a star as a function of the net charge of the droplets of H_2O (above) and DMSO (below). The blue dots are for a positively charged central ion and the red cross for a negatively charged central ion. The dashed lines are a fitting to a quadratic polynomial.

and the charge of a cone is given by a scaling law

$$Q_2 = NC(64\pi^2\gamma\varepsilon_0 H^3)^{1/2} \tag{3.10}$$

where H is the height of a cone, C is a dimensionless proportionality constant, and $64\pi^2\gamma\varepsilon_0 H^3$ has units of charge squared. This form of equation is given in order to make C dimensionless. The scaling law is determined by the fact that the cone is the solution of a minimization of energy, which is the sum of the surface energy term (scaling r^2) and electrostatic energy (scaling Q^2/r). Balancing these two terms leads to that Q^2 is proportional to r^3.

A droplet with three or fewer rays is not described by the model because the surface of the core is not fully covered by the cones.

Even though the water molecules are neutral, the charge in the core region and rays arises from the polarization charge. The conservation of volume of the droplet reads as

$$\frac{4\pi R^3}{3} = N\frac{\pi}{3}H^3\tan^2\frac{\theta}{2} + \frac{4\pi R_1^3}{3} \tag{3.11}$$

where R is the radius of the initial spherical droplet and θ is the half angle of the cone. The area of the core is written as

$$4\pi R_1^2 = N\pi H^2\tan^2\frac{\theta}{2}. \tag{3.12}$$

We solve Equation 3.12 for H and introduce it in Equation 3.11. This yields

$$R^3 = R_1^3\left(1 + \frac{2}{N^{1/2}\tan\frac{\theta}{2}}\right). \tag{3.13}$$

Note that when the number of cones N is large, the radius of the droplet tends to that of the original droplet. This observation is confirmed in the simulations of Ref. [14].

The electrostatic energy of the droplet is written as:

$$E = \frac{1}{2}\frac{1}{4\pi\varepsilon_0}\left[\frac{Q_1^2}{R_1} + \frac{Q_2^2}{R_2} + \frac{2Q_1Q_2}{R_2}\right] = \frac{1}{2}\frac{1}{4\pi\varepsilon_0}\left[Q_1^2\left(\frac{1}{R_1} - \frac{1}{R_2}\right) + \frac{Q^2}{R_2}\right]. \tag{3.14}$$

The surface tension effects are taken into account implicitly by assuming that the system has the conical geometry. In each of the stable conical shapes, the electrostatic pressure exactly balances the surface tension in the system. The surface tension has been taken into account in the expression for Q_2. Therefore, an explicit consideration of the surface energy in Equation 3.14 is expected to have an effect much smaller than the dominant electrostatic contribution. Evidence of that is provided by the simulations, where we observe the inter-conversions between shapes that differ by one ray through ray division or merging.

Equation 3.14 is minimized with respect to R_1 and solved for Q to yield for the lowest power of N

$$Q = 2\cot\frac{\theta}{2}N^{1/2}C(64\pi^2\gamma\varepsilon_0 R^3)^{1/2} + O(\sqrt{N}). \tag{3.15}$$

Solving Equation 3.15 for N entails proportionality between the number rays and the Rayleigh fissility parameter

$$N \approx \frac{\tan^2\theta/2}{4C^2}X. \tag{3.16}$$

For small values of N, we expect this law would break down as the assumption (3.12) does not hold. For the intermediate values of N, the behaviour is qualitatively correct in relation to Figure 3.5.

To investigate the scalability of the star morphologies, we examined different combinations of excess charge amounts and water droplet sizes, maintaining the fissility parameter within the same range ($X = 1.6-1.8$). Table 3.2 clearly shows that four spikes do form regardless of the size of the droplets. However, for the small numbers of H_2O molecules (*i.e.*, N_{H_2O} = 400-500), the rays are not clearly distinguisable from thermal fluctuations of the droplet surfaces, and the rays can be split further to give birth to several additional transient spikes. This is possibly because the volume effect of the central macroion becomes less negligible when the droplets become smaller. Also, based on the scalability of the star morphologies and the experimental studies of Taylor [2], it is expected that the star morphologies will appear in all sizes of the droplets spanning from microscopic to macroscopic ranges.

We emphasize that the fine details of the central macroion structure does not affect the relation between the number of rays and the net charge of the droplet (Equation 3.16). The molecular simulations of a double-stranded DNA [28], a polyhistidine [14], and a protein complex [29, 30] (Chapter 4) all reveal that the formation of star morphologies is a general finding that is independent of the type of macroions.

Table 3.2: Scalability of the star morphologies of charged aqueous droplets of 400-5000 H_2O molecules. N_{H_2O} is the number of water molecules, Q is the amount of the droplet charge, Q_R is the Rayleigh limit, and N_{rays} is the number of rays.

N_{H_2O}	Q (e)	Q_R (e)	Q/Q_R	N_{rays}
400	9	5.58	1.61290	4-5
400	10	5.58	1.79211	4-6
500	10	6.24	1.60256	4
500	11	6.24	1.76282	4-5
1000	15	8.83	1.69875	4
1500	18	10.81	1.66512	4
2000	21	12.48	1.68269	4
2500	24	13.95	1.72043	4
3000	26	15.29	1.70046	4
4000	30	17.65	1.69972	4
5000	34	19.73	1.72326	4

Now we will examine the similarity of the conical shapes that we find in our simulations with those that have been studied by Taylor [2] and later by many other scientists [13, 31] using analytical theory and numerical solutions. The theoretical and numerical studies of the conical shapes of dielectrics determine how the angle of the cone changes with the dielectric constant of the fluid with the assumption of an already existing cone, *i.e.*, the conical shape is assumed in their studies. In the simulations the conical shapes are formed naturally. The angle of the tips is found to be approximately 30° for the aqueous droplets, regardless of the size of the system in the range of charge states that we have investigated. This finding is not in agreement with the estimated Taylor cone angle, which is $\approx 100°$. The change in the angle of the conical shape as a function of the dielectric constant of the fluid is the key quantity in analytical theory. The relative permittivity of a dielectric i (ε_i) is related to the the half angle of the cone θ_0 by the following expression [31] (Equation 1.9):

$$\frac{\varepsilon_i}{\varepsilon_0} = -\frac{P'_{1/2}(\cos\theta_0)P_{1/2}(-\cos\theta_0)}{P'_{1/2}(-\cos\theta_0)P_{1/2}(\cos\theta_0)} \tag{3.17}$$

where $P_{1/2}$ is the Legendre polynomial of the first kind with order $1/2$, and prime denotes derivative. In Figure 3.3.3, we compare the angle estimated for various solvents, including water and DMSO in our simulations, with the analytical findings of Ramos *et al.* [31] from Equation 3.17. Obviously, there is a large difference between the simulation data and the theoretical findings. The value of the dielectric constant of a water droplet is not the reason for this difference since it has been estimated that the dielectric constant in the interior of an aqueous droplet is very close to that of the bulk water [32]. The large difference suggests that the molecular details in the cone region may come into play in determining the cone angles.

3.3.3 Comparison with the Thomson Model

The star formation is reminiscent of the Thomson problem, first proposed by J. J. Thomson [33], which originated from the plum pudding model as a representation of an atomic structure. Although his model was disproved by the discovery of atomic nuclei in the Geiger-Marsden experiment, the Thomson problem still remains significant because it has a plethora of applications in mathematics and science; it captures the competition between local order for neighbouring particles and long-range constraints due to the curvature and geometry [34]. The Thomson problem attempts to determine the stable equilibrium arrangement of N classical electrons constrained on the surface of a sphere and interacting one another via an inverse power law ($F \sim 1/r^x$). In other words, it is to find the minimum energy for a system composed of N identical charges in a sphere subject to spherical constraints $r_i^2 = C$ (where r_i is the distance between charge i and the centre of the sphere and C is a constant equal to the sphere radius squared). The solutions to the Thomson problem may shed light on understanding the electrostatic forces governing systems of different complexities, ranging from atomic electronic structure [35] and multi-electron bubbles in superfluid helium [36] to fullerene patterns for carbon clusters [37], protein arrangements on spherical viruses [38–40], colloidal particles confined at a liquid/liquid interface [41–48], and the surface ordering of metal drops in ion traps [49]. The symmetry and the energetics of the Thomson system may be also found in the distribution of ions in a droplet of solvent with a small dielectric constant especially at lower temperatures where thermal fluctuations are less important than the dominant electrostatic interactions (*i.e.*, large Bjerrum length) [50]. Solvent protrusions found in the star formation resemble the Thomson problem in a sense that the vertices of the rays may correspond to those of polyhedral shapes with symmetry. The distribution of the rays is determined primarily by the equilibrium energetics of the excess charge in a droplet. Also, the excess charge resides

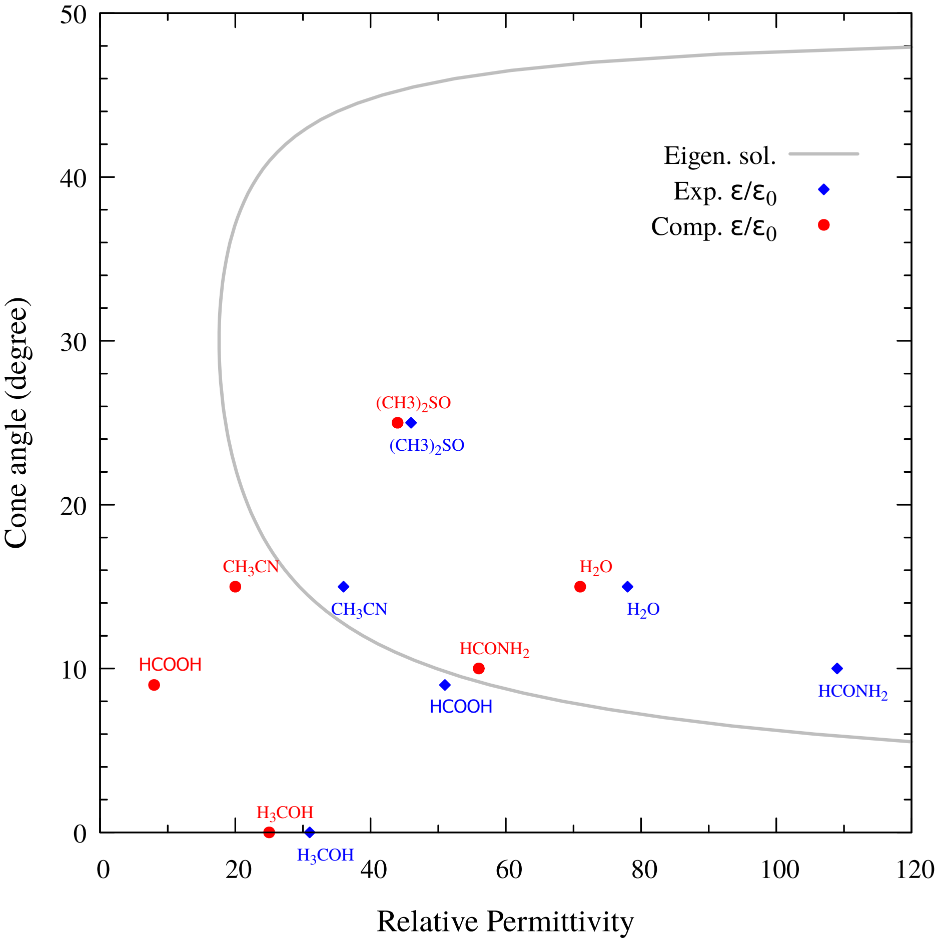


Figure 3.6: Half conical angle *vs*. dielectric permittivity. The gray line is estimated by theory [31] using Equation 3.17. The blue and red squares are the dielectric permittivities from experiments and simulations [25].

on the droplet surface via strong polarization of the polar solvent molecules in response to the electric field emanating from the central macroion. The main difference is that the constraint to the charges residing on a sphere in the Thomson model does not apply to the dielectric droplets. This is because the droplets are deformable in shape due to their surface tension.

3.4 Conclusion

A dielectric spherical nanodrop containing a single highly charged central macroion acquires a "star" shape when the droplet charge-squared-to-volume ratio exceeds a threshold value. We compare droplets composed of water and DMSO which are two dramatically different solvents with respect to their ability to form hydrogen bonds and their molecular size. We found that the same charge-squared-to-volume ratio in the two solvents lead to slightly different numbers of spikes. There are indications that over a wide range of the droplet charges, certain numbers of points are preferred by the droplets. In particular, it is more pronounced in the DMSO droplets which show strong preference to a tetrahedral shape. Also we found that both positively and negatively charged macroions cause the same number of rays in the star formation for water and DMSO droplets if they contain the same magnitude of net charge. Lastly, we presented an analytical description of the number of spikes as a funtion of the droplet charge that is in close agreement with the simulation results. The analysis of cone angles of "star"-shaped droplets of water and DMSO suggested that the molecular details in the cone region, as well as the dielectric constant of the solvents, may come into play in determining the cone angles.

In experiments, the formation of star shapes would be suppressed if dielectric breakdown of the solvent occurs near the highly charged central ion. Computations have shown that water molecules may start to dissociate for electric field strengths around 0.25 V/Å [51]. In the study in Chapter 4 [29], for example, we investigated aqueous droplets with protein complex charge carriers at the charge state +14 e and +17 e. We found that even proteins below their maximum charge state may cause star formations in droplets of certain size. By using dielectric constant value of 4 for water in proximity to a charged macroion [52], the field strength caused by charge +14 e is 0.13 V/Å at 2 mn (where 2 nm is approximately the radius of the proteins that we investigated), which is below the calculated value for water dissociation. We note that we have found that the star shapes are formed with different organic solvents that have higher dielectric breakdown limits. It is therefore likely that dielectric breakdown will not inhibit star formation in experiments.

For generating star-shaped droplets in experiments, the following approaches may be considered: (1) Deposition of solvent molecules on a fixed charged tip or an array of hemispherical caps is a route to be explored in experiments for the creation of the “star”-shaped droplets. (2) Electrospray-generated charged droplets may be used to create “thorny” structures. An electrospray experiment operates by spraying a bulk solution often containing charged proteins or other macroions. The sprayed droplets are charged. Complex processes during the drying-out of the droplets may lead to highly charged macroions. A complete drying-out of the macroion may be needed because chemical reactions even in the latest stage of the macroion desolvation process may increase its charge state. Then, the macroion may be transferred into a chamber of a certain vapor pressure where the solvent may condense on the macroion. Ion-mobility experiments of these highly charged droplets may be used to detect the increased hydrodynamic radius relative to that of a spherical shape. (3) Another possible way for generating star shapes is the usage of a patch clamping experiment [53]. One may use a micro-pipette with an embedded electrode and pipette solution, which in this case can be oil. Calculations show that a dielectric drop that has a high dielectric constant is stable when an electric potential is applied. In solution, one may need to have an oil droplet with dielectric constant greater than 15 immersed in water. A micro-pipette is attractive to use as it would allow one to change the droplet size and the pressure inside the drop. We think that in patch clamping, we do not need instruments related to cell biology.

References

[1] Rayleigh, L. XX. on the equilibrium of liquid conducting masses charged with electricity. *Philos. Mag.* **14**, 184–186 (1882).

[2] Taylor, G. Disintegration of water drops in an electric field. *Proc. Royal Soc. A* **280**, 383–397 (1964).

[3] Sherwood, J. D. The deformation of a fluid drop in an electric field: a slender-body analysis. *J. Phys. A: Math. Gen.* **24**, 4047 (1991).

[4] Miksis, M. J. Shape of a drop in an electric field. *Phys. Fluids* **24**, 1967–1972 (1981).

[5] Luedtke, W. D. *et al.* Nanojets, electrospray, and ion field evaporation: Molecular dynamics simulations and laboratory experiments. *J. Phys. Chem. A* **112**, 9628–9649 (2008).

[6] Duft, D., Achtzehn, T., Muller, R., Huber, B. A. & Leisner, T. Coulomb fission: Rayleigh jets from levitated microdroplets. *Nature* **421**, 128–128 (2003).

[7] Brosseau, Q. & Vlahovska, P. M. Streaming from the equator of a drop in an external electric field. *Phys. Rev. Lett.* **119**, 034501 (2017).

[8] Zeleny, J. Instability of electrified liquid surfaces. *Phys. Rev.* **10**, 1 (1917).

[9] Almería, B., Deng, W., Fahmy, T. M. & Gomez, A. Controlling the morphology of electrospray-generated PLGA microparticles for drug delivery. *J. Colloid Interface Sci.* **343**, 125–133 (2010).

[10] Almería, B. & Gomez, A. Electrospray synthesis of monodisperse polymer particles in a broad (60nm-2μm) diameter range: guiding principles and formulation recipes. *J. Colloid Interface Sci.* **417**, 121–130 (2014).

[11] Malik, S. A. *et al.* Electrospray synthesis and properties of hierarchically structured PLGA TIPS microspheres for use as controlled release technologies. *J. Colloid Interface Sci.* **467**, 220–229 (2016).

[12] Torza, S., Cox, R. G. & Mason, S. G. Electrohydrodynamic deformation and burst of liquid drops. *Philos. Trans. Royal Soc. A* **269**, 295–319 (1971).

[13] Stone, H. A., Lister, J. R. & Brenner, M. P. Drops with conical ends in electric and magnetic fields. *Proc. Royal Soc. A* **455**, 329–347 (1999).

[14] Consta, S. Manifestation of Rayleigh instability in droplets containing multiply charged macroions. *J. Phys. Chem. B* **114**, 5263–5268 (2010).

[15] Consta, S. & Malevanets, A. Disintegration mechanisms of charged nanodroplets: novel systems for applying methods of activated processes. *Mol. Simul.* **41**, 73–85 (2015).

[16] Jackson, J. D. (J. Wiley & Sons, New York, 1998), 3rd edn.

[17] Oh, M. I., Malevanets, A., Paliy, M., Frenkel, D. & Consta, S. When droplets become stars: charged dielectric droplets beyond the Rayleigh limit. *Soft Matter* **13**, 8781–8795 (2017).

[18] Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comp. Chem.* **26**, 1781–1802 (2005).

[19] Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comp. Chem.* **31**, 671–690 (2010).

[20] Yu, W., He, X., Vanommeslaeghe, K. & MacKerell, A. D. Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations. *J. Comp. Chem.* **33**, 2451–2468 (2012).

[21] Koopman, E. A. & Lowe, C. P. Advantages of a Lowe-Andersen thermostat in molecular dynamics simulations. *J. Chem. Phys.* **124**, 204103 (2006).

[22] Jorgensen, W. L. & Jenson, C. Temperature dependence of TIP3P, SPC, and TIP4P water from NPT Monte Carlo simulations: Seeking temperatures of maximum density. *J. Comp. Chem.* **19**, 1179–1186 (1998).

[23] Vega, C. & de Miguel, E. Surface tension of the most popular models of water by using the test-area simulation method. *J. Chem. Phys.* **126**, 154707 (2007).

[24] Chen, F. & Smith, P. E. Simulated surface tensions of common water models. *J. Chem. Phys.* **126**, 221101 (2007).

[25] Caleman, C. *et al.* Force field benchmark of organic liquids: Density, enthalpy of vaporization, heat capacities, surface tension, isothermal compressibility, volumetric expansion coefficient, and dielectric constant. *J. Chem. Theory Comput.* **8**, 61–74 (2012).

[26] Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **14**, 33–38 (1996).

[27] Sharawy, M. & Consta, S. How do non-covalent complexes dissociate in droplets? A case study of the desolvation of dsDNA from a charged aqueous nanodrop. *Phys. Chem. Chem. Phys.* **17**, 25550–25562 (2015).

[28] Sharawy, M. & Consta, S. Characterization of ”star” droplet morphologies induced by charged macromolecules. *J. Phys. Chem. A* **120**, 8871–8880 (2016).

[29] Oh, M. I. & Consta, S. Stability of a transient protein complex in a charged aqueous droplet with variable pH. *J. Phys. Chem. Lett.* **8**, 80–85 (2017).

[30] Oh, M. I. & Consta, S. What factors determine the stability of a weak protein-protein interaction in a charged aqueous droplet? *Phys. Chem. Chem. Phys.* **19**, 31965–31981 (2017).

[31] Ramos, A. & Castellanos, A. Conical points in liquid-liquid interfaces subjected to electric fields. *Phys. Lett. A* **184**, 268–272 (1994).

[32] Simonson, T. Accurate calculation of the dielectric constant of water from simulations of a microscopic droplet in vacuum. *Chem. Phys. Lett.* **250**, 450–454 (1996).

[33] Thomson, J. XXIV. on the structure of the atom: an investigation of the stability and periods of oscillation of a number of corpuscles arranged at equal intervals around the circumference of a circle; with application of the results to the theory of atomic structure. *Philos. Mag. Ser. 6* **7**, 237–265 (1904).

[34] Mehta, D., Chen, J., Chen, D. Z., Kusumaatmaja, H. & Wales, D. J. Kinetic transition networks for the Thomson problem and Smale's seventh problem. *Phys. Rev. Lett.* **117**, 028301 (2016).

[35] LaFave Jr., T. Correspondences between the classical electrostatic Thomson problem and atomic electronic structure. *J. Electrostat.* **71**, 1029–1035 (2013).

[36] Leiderer, P. Ions at helium interfaces. *Z. Phys. B: Condens. Matter* **98**, 303–308 (1995).

[37] Kroto, H. C60: Buckminsterfullerene. *Nature* **318**, 162–163 (1985).

[38] Marzec, C. J. & Day, L. A. Pattern formation in icosahedral virus capsids: the papova viruses and Nudaurelia capensis beta virus. *Biophys. J .* **65**, 2559–2577 (1993).

[39] Zandi, R., Reguera, D., Bruinsma, R. F., Gelbart, W. M. & Rudnick, J. Origin of icosahedral symmetry in viruses. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 15556–15560 (2004).

[40] Bruinsma, R. F., Gelbart, W. M., Reguera, D., Rudnick, J. & Zandi, R. Viral self-assembly as a thermodynamic process. *Phys. Rev. Lett.* **90**, 248101 (2003).

[41] Bausch, A. R. *et al.* Grain boundary scars and spherical crystallography. *Science* **299**, 1716–1718 (2003).

[42] Lipowsky, P., Bowick, M. J., Meinke, J. H., Nelson, D. R. & Bausch, A. R. Direct visualization of dislocation dynamics in grain-boundary scars. *Nat. Mater.* **4**, 407–411 (2005).

[43] Einert, T., Lipowsky, P., Schilling, J., Bowick, M. J. & Bausch, A. R. Grain boundary scars on spherical crystals. *Langmuir* **21**, 12076–12079 (2005).

[44] Kusumaatmaja, H. & Wales, D. J. Defect motifs for constant mean curvature surfaces. *Phys. Rev. Lett.* **110**, 165502 (2013).

[45] Meng, G., Paulose, J., Nelson, D. R. & Manoharan, V. N. Elastic instability of a crystal growing on a curved surface. *Science* **343**, 634–637 (2014).

[46] Irvine, W. T. M., Vitelli, V. & Chaikin, P. M. Pleats in crystals on curved surfaces. *Nature* **468**, 947–951 (2010).

[47] Irvine, W. T. M., Bowick, M. J. & Chaikin, P. M. Fractionalization of interstitials in curved colloidal crystals. *Nat. Mater.* **11**, 948–951 (2012).

[48] Bendito, E., Bowick, M. J., Medina, A. & Yao, Z. Crystalline particle packings on constant mean curvature (delaunay) surfaces. *Phys. Rev. E* **88**, 012405 (2013).

[49] Davis, E. J. A history of single aerosol particle levitation. *Aerosol Sci. Technol.* **26**, 212–254 (1997).

[50] Oh, M. I., Paliy, M. & Consta, S. ”Star” morphologies of charged nanodrops comprised of conformational isomers. *J. Chem. Phys.* **148**, 024307 (2018).

[51] Saitta, A. M., Saija, F. & Giaquinta, P. V. Ab initio molecular dynamics study of dissociation of water under an electric field. *Phys. Rev. Lett.* **108**, 207801 (2012).

[52] Dinpajooh, M. & Matyushov, D. V. Dielectric constant of water in the interface. *J. Chem. Phys.* **145**, 014504 (2016).

[53] Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflügers Arch.* **391**, 85–100 (1981).

Chapter 4

Stability of Weak Protein-Protein Interactions in a Charged Aqueous Droplet

4.1 Introduction

4.1.1 Motivation and Objectives

Protein-protein interactions (PPIs) are orchestrated in an intricate manner to play a pivotal role in all biological processes, including signal transduction, enzymatic catalysis as well as the formation of protein quaternary structures. Elucidating the mechanism of protein complexation and dissociation and finding its kinetics and thermodynamics are therefore important to unveil the fundamental principles underlying biochemical pathways in cells.

To dissect PPIs in a vast array of non-covalent protein complexes, a combination of different experimental techniques have been used in structural biology, which include X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, (cryo-)electron microscopy, computational methods, native mass spectrometry (MS), and isothermal titration calorimetry (ITC) [1, 2]. Among these techniques, native MS is broadly used to find the equilibrium constant of a protein complex [3–13]. Native MS usually operates by preceding electrospray

---

Reproduced by permission of the American Chemical Society from Oh, M. I., Consta, S. "Stability of a transient protein complex in a charged aqueous droplet with variable pH" *J. Phys. Chem. Lett.* **8**, 80–85 (2017).

Reproduced in part by permission of the American Chemical Society from Consta, S., Sharawy, M., Oh, M. I., Malevanets, A. "Advances in modeling the stability of noncovalent complexes in charged droplets with applications in electrospray ionization-MS experiments" *Anal. Chem.* **89**, 8192–8202 (2017).

Reproduced by permission of the Royal Society of Chemistry from Oh, M. I., Consta, S. "What factors determine the stability of a weak protein-protein interaction in a charged aqueous droplet?" *Phys. Chem. Chem. Phys.* **19**, 31965–31981 (2017).

ionization (ESI), since ESI is a soft ionization method.

The detection of protein-protein and protein-ligand dissociation constants (K_d) is in the forefront of the ESI-MS applications [4–6, 14–21]. The reliability of ESI-MS in measuring equilibrium constants of dissociation has been validated by comparisons with other experimental methods such as ITC [5, 17, 22–24]. The major challenges in the measurement of the equilibrium constants by ESI-MS are commonly found in weakly bound protein complexes, which are characterized by their K_d values in the micromolar to millimolar range. Over the last two decades, there has been a substantial improvement on the experimental protocols that aim at maintaining "fragile" protein-protein interactions (PPIs) during the transfer from bulk solution to the gas phase [1, 14, 16, 19, 20, 25–28]. Nonetheless, the reliable detection of the dissociation constant of a class of weak noncovalent protein complexes still remains questionable [14].

Due to the unknown role of droplet chemistry in the stabilization of weak noncovalent interactions, the question on whether the gas-phase ensemble of the complexes is reflective of the chemical equilibria of the species in bulk solution has been debated [3, 5, 6, 13–19, 26]. For instance, Zenobi *et al.* [18] argued that using ESI for analyzing biomolecular complexes is not yet routine, and specific protocols, characteristic of each biomolecular complex, may be required. Particularly, transient and weak noncovalent interactions are prone to be ruined during multistep and/or harsh isolation procedures, as these interactions are often sensitive to many factors such as temperature, acidity, and salt concentration.

The goal of this study is to establish the principles that govern the stability of a weak transient protein complex in an aqueous droplet. In an ESI-MS experiment, each of electrosprayed nanodrops may contain at most one protomer or the dimeric complex. Therefore, one of the manners in which the equilibrium constant can be modified in the droplet is by the dissociation of the complex. This is the mechanism that is examined here in the latest stage of the droplet lifespan.

The specific objectives of the study are: (a) we propose a method that allows for the computation of the correction in the equilibrium constant between that measured in the bulk solution and that measured in the gaseous state, (b) we elucidate the dissociation mechanism of a weakly bound protein complex, (c) we examine the factors that may affect the stability of the weak PPIs, and (d) we propose how to compute the contribution from the untoward complex dissociation due to the droplet environment to the error in the ESI-MS measurement of K_d.

For this study, two weakly bound noncovalent complexes are examined: (a) that of ubiquitin (Ub) and the ubiquitin-associated (UbA) domain from DNA-damage-inducible 1 protein (Ddi1) (RCSB PDB [29, 30] code 2MRO [31]) and (b) that of two Ub molecules forming the closed conformation of a lysine-48-linked diubiquitin complex (RCSB PDB code 2PEA [32]). The former complex is a weak transient protein complex in solution with dissociation equilibrium constant in the micromolar range [31]. Two free ubiquitin molecules are found to dimerize noncovalently in solution with a dissociation constant in the millimolar range [33], forming a (ultra)weak transient protein complex. 2MRO is involved in proteolytic degradation through which many biochemical pathways are regulated at the cellular level. The latter complex (2PEA) is found when ubiquitin molecules are dimerized (or polymerized) for multivalent binding with ubiquitin-binding (UbB) domains by forming an iso-peptide bond at the site of lysine-48.

2MRO and 2PEA are good candidates for the purpose of this research for several reasons. Since 2MRO and 2PEA consist of only 120 and 152 amino acids, respectively, their sizes allow for computational investigation. Since PPIs are highly sensitive to constituent residues (*e.g.*, protonation states and hot spots) and dependent on protomer shapes (*i.e.*, shape complementarity), we selected atomistic modelling to represent these effects better than coarse-grained modelling. In coarse-grained modelling, proteins are often modelled by a sphere with isotropic attraction or patches. This manner of modelling allows for the study of the collective motion of complex PPIs, such as self-assembly of proteins and protein crystallization that take place on large spatial and temporal scales [34–37]. The anisotropic, specific, and globular nature of proteins are well represented by the patchy particles, yet they are not suitable for studying the interaction in a two-component protein complexes. Therefore, the more detailed and computationally expensive atomistic modelling was selected instead of the coarse-grained modelling. Moreover, 2MRO has been well studied experimentally using multidimensional NMR [31], and the noncovalent dimerization of ubiquitin has been investigated using solution NMR techniques and analytical ultracentrifugation analysis [33]. Also, ubiquitin is known for its high thermal and acidic resistance, so it has been widely selected for both experimental and computational studies [38]. Although this study is on specific protein complexes, 2MRO and 2PEA are only used as paradigms to demonstrate methodologies that are are applicable to other protein complexes.

4.1.2 Characterization and Classification of Protein-Protein Interactions

Specific PPIs in solution rest heavily on noncovalent interactions between constituent protomers, including hydrogen bonding and ionic, hydrophobic, van der Waals and π-π interactions. The strength of a typical salt bridge is in the range of 12.5−17 kJ/mol (3−4.1 kcal/mol). Hydrogen bonds are in the range of 2−6 kJ/mol (0.5−1.4 kcal/mol) in water, yet the free energy increases up to 12.5−21 kJ/mol (3−5 kcal/mol) when either the donor or acceptor is charged. Short-range van der Waals interactions are typically 4 kJ/mol (1 kcal/mol), yet the energy increases up to 4−17 kJ/mol (1−4.1 kcal/mol) in the interior of a protein. The average thermal energy at room temperature is 2.5 kJ/mol (0.6 kcal/mol) [39]. Polar weak interactions have variable strengths as long-range electrostatic interactions depend largely on the distance between two charge sources and their chemical environment [39]. Electrostatic interactions may be enhanced in a vacuum and in non-polar regions away from bulk solution, but they become weak in a dielectric solvent such as water. These interactions are vital for proteins recognizing their partners in a highly crowded environment.

PPIs are indeed diverse. These interactions are classified into different categories, depending on the composition and the binding affinity in addition to the lifetime or stability of the protein complex [40]. When PPIs occur between identical protomers for association, the complex is said to be homo-oligometic, whereas the constituents in a hetero-oligomeric complex are asymmetric. When individual protomers are found as free, stable structures on their own in vivo, the complex is said to be non-obligate, whereas subunits that constitute an obligate complex cannot exist free on their own. PPIs are classified as either transient or permanent depending on the lifetime (or stability) of the protein complex. A permanent complex is usually characterized by its high stability and thus existence only in an irreversible complexation. In contrast, a weak transient complex features dynamic association and dissociation at equilibrium in solution, whereas a molecular trigger is required for a strong transient complex to shift the oligomeric state [40]. The former has a dissociation constant (K_d) typically in the micromolar (μM) range, and the latter may have K_d values in the nanomolar (nM) range [40]. Transient PPIs are significant as found in signaling and regulatory mechanisms so that the cell is able to make a swift response to extracellular stimuli [41]. Transient PPIs may be further subdivided into weak and strong interactions. Strong transient interactions, stabilized by binding of an effector molecule, may last longer and have a lower K_d in the nanomolar range. They shift equilibrium of association/dissociation under certain disturbances, whereas weak transient interactions break and form continuously [40]. Other than K_d, transient PPIs are differentiated from permanent PPIs in many other aspects such as lower interface contact area (ΔASA), higher interface polarity, and smaller conformational changes upon binding [41].

4.2 Modelling and Computational Methods

4.2.1 Assignment of the Charge State of a Protein

In this study we treat the stability of a protein complex in an evaporating droplet. This is a challenging problem because the droplet acidity changes as the droplet changes in size. By using the equation of the Rayleigh limit (Equation 1.6) one finds that the pH of a water droplet decreases by 1.5 when its radius is reduced by an order of magnitude [42]. In this section we describe how to assign the charge state of a protein in a droplet of a specific size, which is at equilibrium. In the next section we shall discuss how we treat the complex stability in the constantly changing droplet environment.

We find the protein charge state by using the macroscopic equations for the equilibrium constant and the constraint on the droplet charge state imposed by the Rayleigh limit. It is noted that in a bulk solution, the net charge of the protein is determined simply by the solution pH which is an externally controlled variable. In contrast, the pH of a droplet is determined by the Rayleigh limit [43–45], which is a physical constraint on the maximum charge a droplet can sustain before it becomes unstable.

Let us consider a droplet that is composed of water, a protein, ammonium and acetate ions as well as hydronium and hydroxide ions. The addition of ammonium acetate in the parent solution is a commonly found in ESI-MS experiments. The average total droplet charge is given as the sum of the charges of all the chemical species present in the droplet. For example, if the free charge carriers are ammonium and acetate ions, the total droplet charge is given by the following expression:

$$Q_R = Q_p(\mathrm{pH}) + N_A \cdot V \cdot 1000 \cdot \left([\mathrm{NH_4}^+] - [\mathrm{CH_3COO^-}] - [\mathrm{OH^-}] + [\mathrm{H^+}]\right) \tag{4.1}$$

where Q_R is the droplet charge at the Rayleigh limit (*i.e.*, $X = 1$ in Equation 1.6), N_A is Avogadro's number, V is the droplet volume, Q_p is the protein charge at a given pH, and the square brackets denote the molarities of the components. Equation 4.1 holds on the average as an ensemble of droplets is considered. MD simulations have demonstrated that over the course of the droplet evaporation the charge and the radius of the droplet follow closely the Rayleigh expression given by Equation 1.6 [46]. In the droplet at the latest stage of its evolution, the volume of solvent becomes zero. Hence, for the model defined by Equation 4.1, the final charge state of a desolvated protein is always determined by the Rayleigh charge of a droplet with the same volume as the volume of the protein.

The species concentrations are not independent but they are calculated from their chemical equilibria.

$$[\mathrm{OH^-}]_a \cdot [\mathrm{H^+}]_a = K_w \tag{4.2}$$

$$10^{-\mathrm{pH}} = [\mathrm{H^+}]_a \tag{4.3}$$

$$\frac{[\mathrm{NH_4}^+]_a}{[\mathrm{H^+}]_a \cdot [\mathrm{NH_3}]} = K_{\mathrm{NH_4^+}} \tag{4.4}$$

$$\frac{[\mathrm{CH_3COOH}]}{[\mathrm{CH_3COO^-}]_a \cdot [\mathrm{H^+}]_a} = K_{\mathrm{CH_3COO^-}} \tag{4.5}$$

where K_w is the equilibrium constant for water autoprotolysis at temperature T, and the logarithm of the molarity of the hydronium ions is equated to the apparent pH. The subscript a denotes activities of the ionic species. We will be mixing concentrations and activities in the discussion that follows. In the numerical calculations and the plots presented in the discussion, the Debye-Hückel correction was used for the activity coefficient for a point charge [47]. The activity coefficient ζ_i for an ion i in this approximation is given by $\log \zeta_i = -C z_i^2 I^{1/2}$, where z_i

is the charge of the ion i and I is the ionic strength of the solution, and C is a constant. For hydronium and hydroxide ions, we use concentrations instead of the activities. In Equations 4.5, the activity coefficients are transferred on the equilibrium constants (right-hand side of the equations), and thus, we are left with the concentrations in the ratios. The aim is now to solve numerically the system of Equations 4.1-4.4 coupled to the Rayleigh criterion (Equation 1.6). To this end, an expression of the charge on a protein Q_p(pH) is derived.

It is obvious that the average net charge of a protein is pH-dependent. The given charge state of a protein comprises all possible protonation substates where the charge is distributed among the titratable groups of the protein. The relative sampling weight of each substate is determined by the local electric potential as well as the protein conformation ensemble of the substate. With some degree of scientific confidence, this can only be achieved by using quantum mechanics/molecular mechanics (QM/MM) methods where one allows transfer of a proton between different ionizable groups. The intrinsic pK$_a$ value of an ionizable group differs from that of an amino acid in solution, as it is shifted by an environmental perturbation in a protein such as desolvation effects and changes in interactions with neighbours. Several methods have been developed to assign the charge state of each ionizable group based either on sequence similarities, protein structures, or atomistic modelling [48]. However, accurate calculations of their pK$_a$ values still remain highly challenging in structural biochemistry [49–59]. Therefore, a direct sampling of various protein charge states is rather impractical. In the presented method, one can find the average charge of the protein using the macroscopic equations and then assign the charge in the protein by first protonating the residues with the highest pK$_a$ values. If the average charge is fractional, one has to consider different charge states around the value. Also, it should be reminded that the pK$_a$ values of the ionizable groups differ from those of the native amino acids in solution and thus can have a range of values depending on their local chemical environments. As a result, when the pK$_a$ values of some amino acids are quite close to one another, one can expect the presence of different protonation states for the same pH value. Therefore, if the pK$_a$ values of different ionizable groups are similar, the role of different charge distributions in complex dissociation must be examined.

Nevertheless, one can still find the expectation value of the total charge by solving the system of chemical equilibrium equations with the equilibrium constants determined by the pK$_a$

values. The method for calculating the total average charge on a protein at a given pH is as follows. In our model, it is assumed that the number of charged groups is large and the distribution of pK_a obtained from experimentally obtained data set is normally distributed. An average titration curve is calculated for a typical protein by replacing the values of the individual pK_a with the corresponding distribution. As an example, we consider histidine (His), a weak basic amino acid. The pK_a of the corresponding amino acid is 6.5 [60]. Knowing the pK_a of the amino acid in a protein, from the definition of pK_a, one can calculate the relative population $C(\mathrm{pK}_a, \mathrm{pH})$ of the ionizable group His^+ in the protonated state by using the Henderson-Hasselbalch equation:

$$\frac{[\mathrm{His}^+]}{[\mathrm{His}]} = 10^{\mathrm{pK}_a - \mathrm{pH}}$$

Therefore,

$$C(\mathrm{pK}_a, \mathrm{pH}) = \frac{[\mathrm{His}^+]}{[\mathrm{His}^+] + [\mathrm{His}]} = \frac{1}{1 + 10^{\mathrm{pH} - \mathrm{pK}_a}} \tag{4.6}$$

In a protein, the pK_a values of any ionizable group depends on its local environment. To find the expectation value of the charge of the same type of amino acids, all the charges must be summed up:

$$Q = \sum_i C(\mathrm{pK}_{a_i}, \mathrm{pH}) \tag{4.7}$$

where the pK_{a_i} values are unknown. In this method, the pK_{a_i} values are taken from experimental data [60]. We use $\mathrm{pK}_a = 6.6 \pm 1.0$ and replace a set of values for pK_a for all the residues of the same type with the normal distribution

$$P(\mathrm{pK}_a) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{[\mathrm{pK}_a - \langle \mathrm{pK}_a \rangle]^2}{2\sigma}} \tag{4.8}$$

where $\sigma = \Delta \mathrm{pK}_a$ is the width of the distribution of the experimental pK_a values. Introducing Equation 4.8 into Equation 4.7, the total charge expectancy for N_{Aag} number of basic residues of type Aag (by Aag we denote an amino acid in general) is given by the following integral:

$$Q_{\mathrm{Aag}}(\mathrm{pH}) = N_{\mathrm{Aag}} \int_{-\infty}^{\infty} P(\mathrm{pK}_a) C(\mathrm{pK}_a, \mathrm{pH}) d\mathrm{pK}_a. \tag{4.9}$$

For the various pH values, numerical integration of Equation 4.8 is performed. $Q_{\mathrm{Aag}}(\mathrm{pH})$ as a function of pH is then fitted by a nonlinear fitting procedure to the following function

$$\hat{Q}(\mathrm{pH}) = \frac{N_{\mathrm{Aag}}}{1 + a^{\mathrm{pH} - \langle \mathrm{pK}_a \rangle}} \tag{4.10}$$

where a (fitting parameter) and $\langle \mathrm{pK}_a \rangle$ vary for different ionizable groups. Calculations show that across all values of pH, the difference between the fit and the integral (Equation 4.9) does not exceed 0.01 (Figure 4.1).

For the acidic residues, Equations 4.6 and 4.10 are replaced by the corresponding equations

$$\begin{aligned} C(\mathrm{pK}_a, \mathrm{pH}) &= -\frac{1}{1 + 10^{\mathrm{pK}_a - \mathrm{pH}}} \\ \hat{Q}(\mathrm{pH}) &= -\frac{N_{\mathrm{Aag}}}{1 + a^{\langle \mathrm{pK}_a \rangle - \mathrm{pH}}} \end{aligned} \tag{4.11}$$

The charge of the overall protein (denoted by Q_p) is the sum of the charges found for the residues at a specific pH value. Therefore, Q_p is a function of pH. In Equation 4.1, Q_p and all the concentrations (from Equations 4.2-4.5) are expressed as a function of pH. Equation 4.1 is solved iteratively, and it provides the pH of a droplet of a certain size. Once the pH is known, Q_p can be computed. Figure 4.2 shows the titration curve for 2MRO protein complex. Over the range of pH from 0 to 14, the protein charge changes monotonously from $+18\ e$ to $-17\ e$ (where $e = 1.602 \times 10^{-19}$ C is the elementary charge). Using the derived expression for $Q_p(\mathrm{pH})$, one solves the system of Equations 4.1-4.5 coupled with the Rayleigh criterion (Equation 1.6). The solution of the system of equations provides the charge of the protein that it is subsequently used to study the dissociation of the protein complex.

4.2.2 Method for the Computation of the Rate of Complex Dissociation

In order to study the effect of the droplet environment in the complex dissociation, it is ideal to perform non-equilibrium MD runs starting the simulations from a large number of different initial conditions of droplets (in a similar way that an experiment operates). It is emphasized that hat during these runs, the droplet pH should change constantly, and therefore, the protonation state of the protein may change. The pH of a particular droplet size with an embedded protein

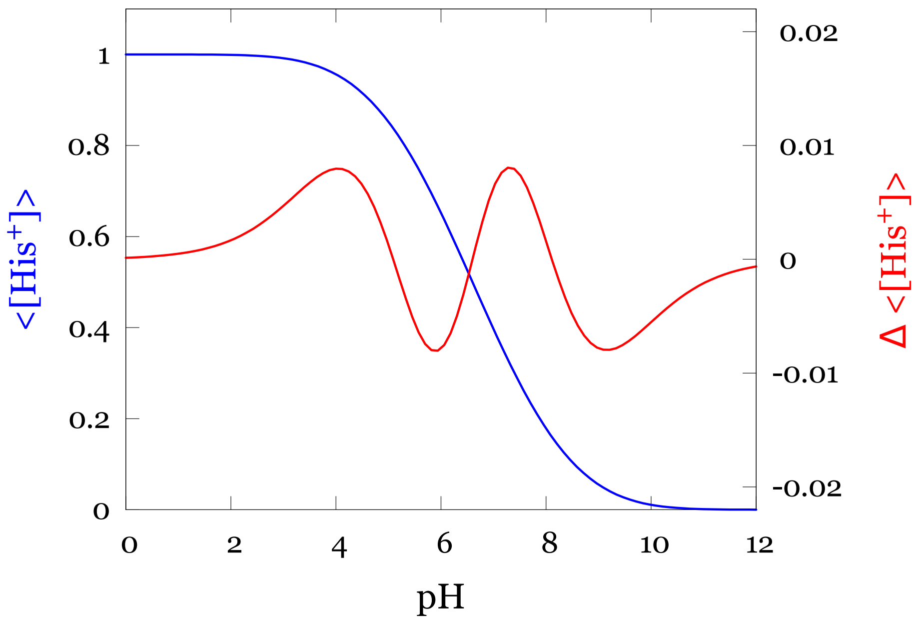


Figure 4.1: (a) The blue line is the titration, and (b) the red line is the difference between the charge estimated by the integral (Equation 4.9) and the charge from the fit (Equation 4.10). The two curves use the different y-axes scale to highlight the difference in magnitude.

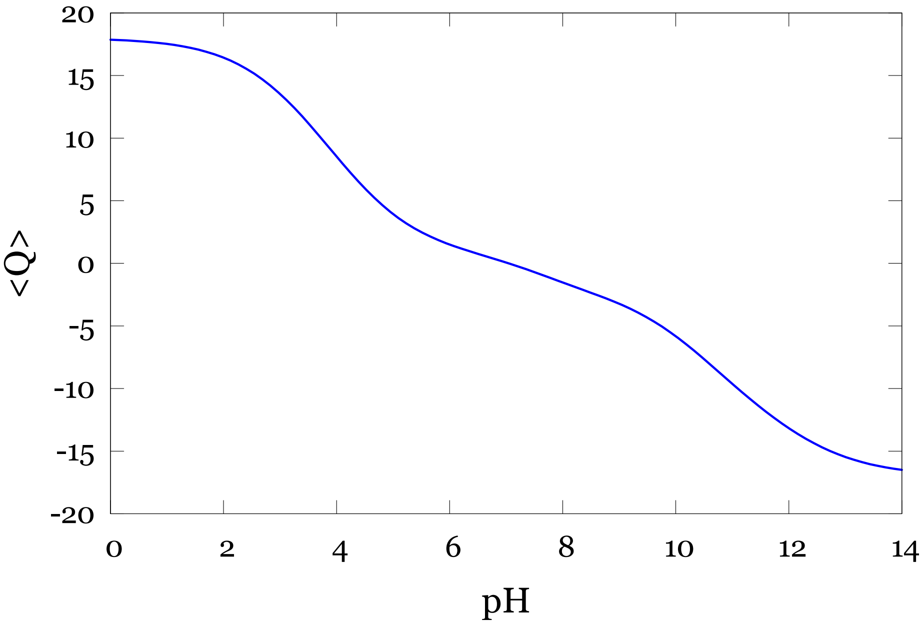


Figure 4.2: $\langle Q \rangle$ for the entire 2MRO from the fit *vs*. pH.

is already described in Section 4.2.1. In the "ideal" runs proton transfer reactions should take place in the course of droplet evaporation. The proton transfer reactions involve the breaking and formation of covalent bonds, therefore, they have to be treated by quantum chemistry methods. However, MD calculations using quantum chemistry modelling are not feasible in practice for the droplet sizes under investigation.

As the non-equilibrium runs with variable pH are rather non-feasible, the following method was devised: We first perform equilibrium MD runs at constant number of molecules (N), volume (V) and temperature (T) (NVT or canonical ensemble) and constant-N, V and energy (E) (NVE or microcanonical ensemble) by enclosing the entire protein complex-H_2O droplet in a spherical cavity (Figure 4.3). During the equilibration state of the simulations part of the solvent evaporates from the droplet and an equilibrium vapour pressure is established. The combination of the cavity volume and the temperature determines the average size of the connected droplet. The conditions in the cavity does not intend to mimic the conditions in an ESI-MS experiment.

The complex dissociation rate should be computed for different droplet sizes that have different droplet pH values and so protein protonation states. Therefore, in the computation of the dissociation rate, MD trajectories starting from a number of initial conditions are generated up to a cut-off time. The process relevant to the ESI-MS experiment is the competition between the rate of droplet evaporation and the rate of complex dissociation. Therefore, the cut-off time of the MD runs in the cavity is determined by performing non-equilibrium MD runs of the droplet solvent evaporation *in vacuo*, as the droplet follows non-equilibrium or quasi-equilibrium evaporation in experimental conditions. This is an independent estimate and depending on the specific ESI conditions, one may compute the solvent evaporation rate under different conditions. Specifically in our simulations, the cut-off time of the MD runs in the cavity is set to be the required time for the droplet to reduce the number of the solvent molecules to a few hundred of H_2O molecules. In this study, based on the non-equilibrium runs, the cut-off time was set to 10 ns and ≥ 6 ns for the droplet systems of 2MRO at 370-390 K and 2PEA at 360-390 K, respectively.

Here we emphasize again that the simulations of the complex dissociation in a droplet should not be viewed as direct mimicry of an ESI process. The comparison between the com-

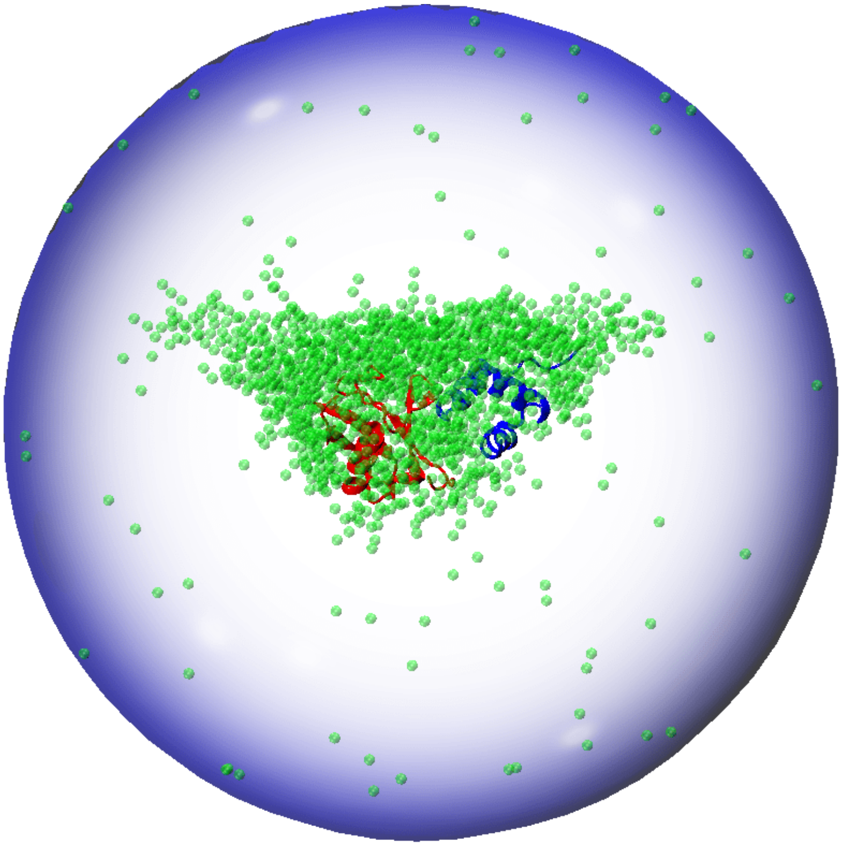


Figure 4.3: Computational set-up for 2MRO with +14 *e* in a droplet of 2000 H_2O molecules. The protein complex is coloured in red (ubiquitin) and blue (ubiquitin-associated domain), and only the oxygen sites of the water molecules are shown in green for clarity. In order to perform equilibrium simulations, the entire droplet system is enclosed in a cavity (coloured in transparent blue). The radius of the cavity used in this study is 20 nm.

plex dissociation rate and the rates of other relevant processes in the experiment (such as droplet evaporation) will provide information on whether the complex dissociation in the droplet will affect the equilibrium constant measurement using ESI-MS. If the solvent evaporation rate is higher than the complex dissociation rate, then the complex is likely to stay bound until the droplet reaches the next distinct smaller size in which the solvent acidity will be different, and therefore, the protonation state of the protein will be altered. (We remind that the protonation state of the protein complex is determined according to the discussion in Section 4.2.1.) For the smaller droplet size, one needs to repeat the rate calculations in the cavity, and this procedure is repeated for even smaller droplet sizes. Conversely, if the complex dissociation rate is higher than the droplet evaporation rate, then the populations of the bound and unbound states of the complex would change in the gaseous state relative to those in bulk solution (see also Section 4.3.6). Here, simulation data are presented only for several specific droplet sizes in order to demonstrate how to apply the method.

4.2.3 Simulation Details

Atomistic molecular dynamics (MD) simulations were performed for two protein complexes: (a) a complex of ubiquitin (Ub) and the UbA domain from DNA-damage-inducible 1 protein (RCSB PDB [29, 30] code 2MRO [31]) and (b) a complex of two Ub molecules forming a lysine-48-linked diubiquitin (RCSB PDB code 2PEA [32]). All simulations were performed by using the NAMD package version 2.9 [61]. Both complexes were studied in an aqueous droplet environment. For equilibrium simulations, every droplet system was confined in a spherical cavity of radius 20 nm. The cavity was generated by applying the spherical boundary condition. The spherical cavity had a radius of $r \approx 20$ nm (Figure 4.3) generated by applying the spherical boundary condition. The cavity is sufficiently large to accommodate even the largest droplet shape fluctuations. Both electrostatic and van der Waals interactions were treated with a switch scheme at the boundaries of the cavity. Each droplet system was then thermalized at 390 K for 1 ns by using the Lowe-Andersen thermostat [62]. After equilibration, the thermostat was turned off so as constant-energy dynamics was performed. The protein complexes were modelled using the CHARMM General Force Field [63, 64], and the H_2O molecules were modelled by using the TIP3P model [65]. MD was performed using the velocity-Verlet algorithm with a time step of 1.0 fs. For making estimates using the Rayleigh limit expression, the surface tension and the density values of the TIP3P water model were obtained from Refs. [66–69]. VMD version 1.9.2 [70] was used for visualization purposes.

Charged 2MRO in Aqueous Droplets

Simulations were performed for 14 droplets comprised 2000 H_2O molecules and 2MRO in charge state +14 *e*. The temperature of the systems was in the range of 370−390 K. The charge +14 *e* is distributed as +11 *e* and +3 *e* to Ub and the UbA domain, respectively. Different charge distributions of the charge +14 *e* to the protomers are also presented in Section 4.3.3, but we expect that these charge distributions make a smaller contribution to the entire droplet ensemble. After relaxation, equilibrium between water vapour and the droplet is established within the cavity. Evaporated water molecules fill the vacant space in the cavity, leaving ≈ 85 % and ≈ 70 % of the initial droplet mass at 370 K and 390 K, respectively.

The study of the complex dissociation includes various parameters to examine with respect to the charge state of the protein complex and the size of the droplet. In this study, the system of interest is found in the late stage of the droplet lifetime in an ESI process. We performed the majority of the 2MRO simulations in droplets composed of ≈ 1350−1650 H_2O molecules and the complex ion charged at +14 *e*. Figure 4.4 (a) shows that 2MRO may hold a net charge of +14 *e* in a droplet containing a few thousands of H_2O molecules when the concentration of ammonium acetate is low. For comparison, two additional simulations were also performed for droplets consisting of 1250−1400 H_2O molecules and 2MRO with zero and +10 *e* net charges in the main body for 10 ns following 1 ns of equilibration.

Here we selected stringent conditions by preparing spherical droplets that are found above the Rayleigh limit (therefore they are unstable) and elevated temperature. An important question that arises is the relation of the simulation conditions used in this study to those in ESI experiments. The temperature of the droplets in ESI experiments is still an open question despite the very insightful experiments that have been performed [71, 72]. The experiments of Antoine *et al.* [71] and Cook *et al.* [72] detect the temperature of droplets with linear dimensions in the micrometer regime. Both studies agree that the nanoscopic droplets for which they do not have measurements the expectation is that they will undergo faster thermalization than the micro-droplets with the warm background gas. Therefore, it is expected that the nanoscopic droplets will be found at elevated temperature relative to room temperature.

To study the effect of temperature on the stability of the protein complex, two droplet systems were simulated at 300 K for 15 ns following 1 ns of equilibration. The systems contained 1400 H_2O molecules and a $2MRO^{14+}$ in the remaining connected droplet after equilibration. This size of the charged droplet is found $\approx$ 30 % above the Rayleigh limit. In one of the simulations, the two protomers were interacting via the stabilizing hydrophobic interface, while in the other, the two protomers were connected by the destabilizing interface (explained in Section 4.3.2) that leads to complex dissociation at 370−390 K.

In order to study the effect of charge distribution on the stability of the protein complex ion two additional sets of two constant-temperature MD simulations were carried out at 390 K. In the one set, the charge is distributed on the $2MRO^{14+}$ complex as +10 e and +4 e on the ubiquitin and the UbA domain, respectively. In the other set of the simulations, the distribution is +12 e on the ubiquitin and +2 e on the UbA domain. In each of the sets, one simulation starts from the $2MRO^{14+}$ complex where the protomers interact with each other via their hydrophobic surfaces, and the other simulation starts from the complex ion with the destabilizing interface as were identified from the simulations of $2MRO^{14+}$ with charge distribution +11 e on the ubiquitin and +3 e on the UbA domain (Section 4.2.3).

Charged 2PEA in Aqueous Droplets

To generalize our findings to other weakly bound protein dimers, we performed 10 equilibrium MD simulations (five NVT and five NVE runs at 360−390 K) of an aqueous droplet composed initially of 833 H_2O molecules and the noncovalent diubiquitin complex at the charge state +14 e. The system was built based on the experimental result that Servage *et al.* [73] obtained: The plots of arrival time distribution versus m/z for ubiquitin ions obtained from cryogenic ion mobility-mass spectrometry (cryo-IM-MS) [73] reveal that the dimer is detected at a charge state +14 e with a high degree of hydration ($N_{H_2O} \approx 285$). The dissociation constant (K_d)

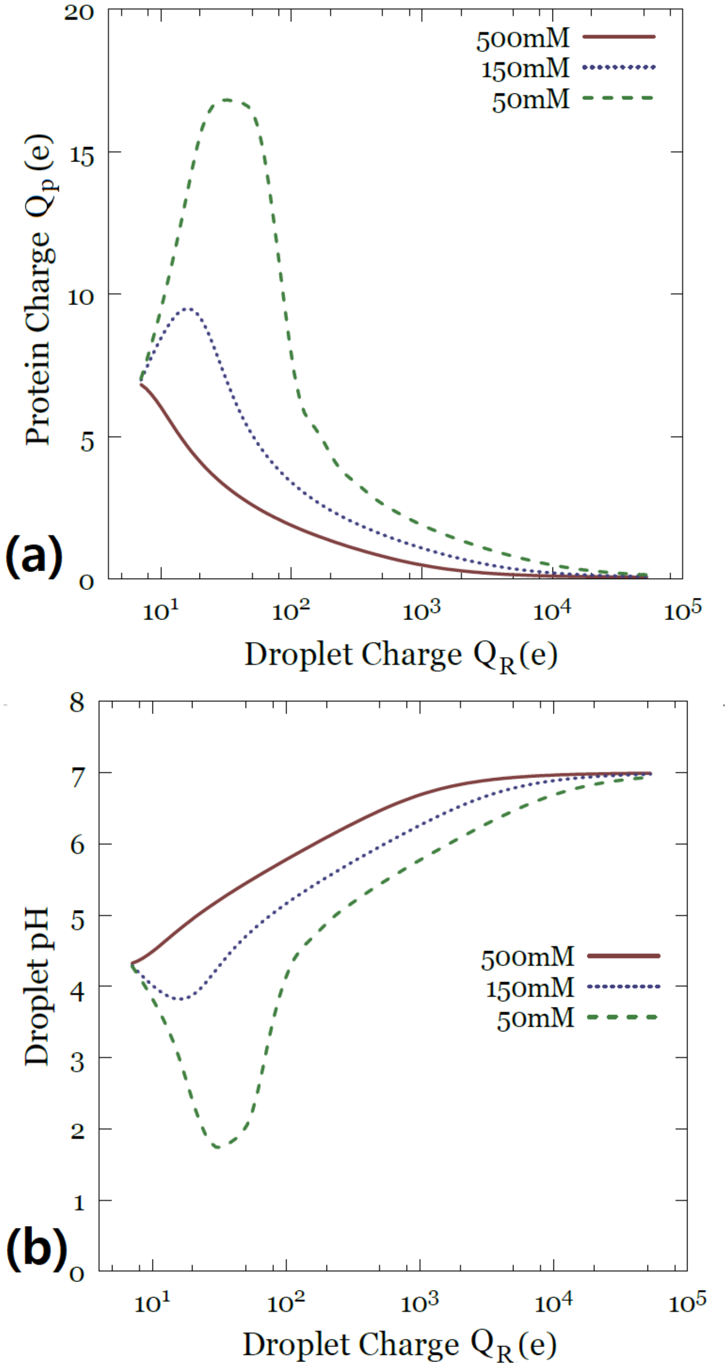


Figure 4.4: (a) Protein charge Q_p (e) *vs*. droplet charge Q_R (e) for various concentrations of CH_3COONH_4. (b) Droplet pH *vs*. droplet charge Q_R (e) for the same concentrations of CH_3COONH_4 as those in (a).

of free ubiquitin molecules dimerizing to a noncovalent dimer in solution has been measured by using NMR spectroscopy [33], which is found to be in the millimolar range. To examine the possible existence of distinct droplet morphologies when a spherical droplet is above the Rayleigh limit the droplet system was also simulated at 300 K for 3 ns. The same simulation protocol was followed as described in Section 4.2.3.

4.3 Results and Discussion

4.3.1 Description of the Ub-UbA interface

A protein-protein interface is of high complexity and specificity and may be characterized by such descriptors as its area, shape, and surface complementarity. When the stability of a noncovalently bound protein complex is concerned, there is an important question to be asked: "What are the residues involved in the formation of the protein-protein interface?". As protein-protein interfaces are often hydrophobic, hydrophobicity is a driving force in the recognition of a partner protein [74]. It is known that the hydrophobic patch centred on isoleucine-44 of ubiquitin serves as the recognition site bound by most UbB domains including those found in a shuttle protein [75]. UbA domains, which form a compact bundle of three α-helices stabilized by a hydrophobic core, have hydrophobic surface patches, and these patches often act as binding sites for other proteins [76]. Consequently, UbA domains have been found to play a crucial role in many other PPIs [77]. Direct visualization of the protein complex reveals that the initial interface formed by ubiquitin and the UbA domain in 2MRO is indeed mostly hydrophobic. Figure 4.5 (a) shows what residues are involved in the formation of the interface in our protein complex. The orange and black dots indicate two amino acids (identified by their residue numbers on axes) are within 8 Å and 4Å in distance, respectively. The distance was measured based on the locations of their β-carbons to account for the orientation of their side chains. Also, glycine was omitted due to the lack of its side chain. The analysis of this contact map reveals that 62 % of the amino acids at the interface is sorted as hydrophobic ones. Interestingly, many of the polar and charged amino acids are located only at the outer verge of the interface with direct access to water. Therefore, the core region of the interface is predominated by the presence of hydrophobic residues, as supported by the computational prediction of hot spots [78, 79].

Hot spots are a small fraction of residues in the interface that cause destabilization of the bound state, leading to an increase in the binding free energy of at least 2.0 $kcal/mol$, upon alanine mutagenesis. They make large contributions to the stability (or lifetime) of the protein complex, and they are present mostly in a central region of the interface, hidden from solvent [80]. As shown in Figure 4.5 (b), the hot spots are leucine-8, isoleucine-44, valine-70, and leucine-73 in Ub, and alanine-423, serine-424, and phenylalanine-427 in the UbA domain. Except serine-424, the rest of the hot spots are hydrophobic residues.

The buried surface area (BSA) of a protein complex is a hydrophobic surface removed from contact with solvent and buried in the interface. BSA is a descriptor related to the binding affinity of a PPI, whose magnitude is estimated to be approximately 0.025 kcal/mol per 1 $Å^2$ of the hydrophobic surface. It was calculated that our protein dimer has $\approx$ 624 $Å^2$ which is in the range of a typical transient protein complex [41]. Based on the estimation, the binding free energy at the interface is estimated to be $\approx$ 15.6 kcal/mol. The dissociation constant (K_d) of the complex of monomeric Ub and the UbA domain has been measured to be 150 ± 16 mM by analyzing NMR titration curves [31], which is again in the typical range of transient PPIs.

4.3.2 Dynamics and Mechanism of 2MRO Dissociation

Dissociation of the complex was found in eight out of the 14 simulation runs. If occurs, the complex dissociation is always followed by division of the droplet within 10 ns. Figure 4.6 (a) shows the times required for the complex separation and droplet fission. The simulations reveal that the time required for the protein complex to dissociate, when it occurs, varies between 4 to 6 ns. The droplet fission follows within 300 ps to 1.3 ns. It is important to note here that the complex dissociation may occur much earlier than 4 ns, depending on the state of the complex conformation in the parent bulk solution.

To find out whether the complex will dissociate during the droplet lifetime, we have to perform non-equilibrium evaporation runs of the droplet and compare the rate of complex dissociation as was determined by the equilibrium runs (Figure 4.6 (a)) to the rate of solvent

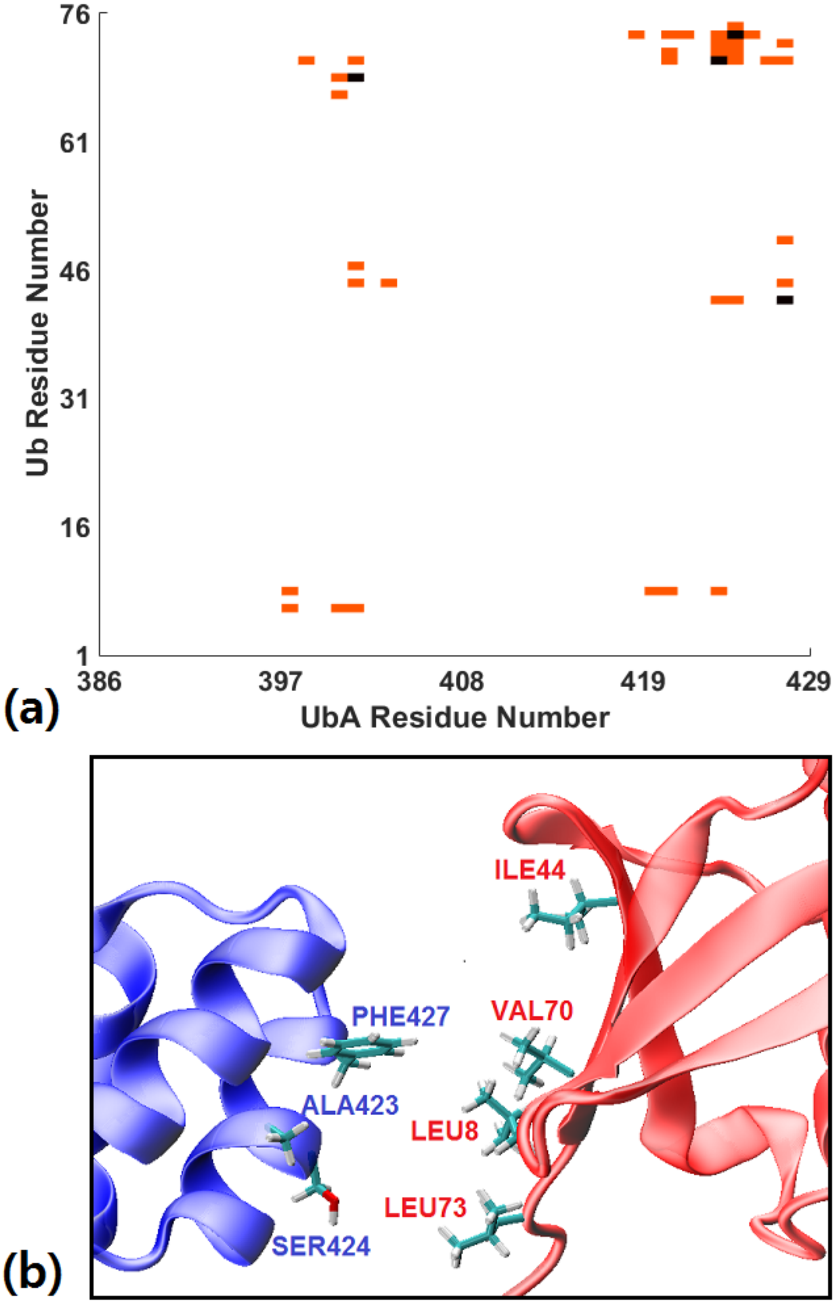


Figure 4.5: (a) Contact map showing residues involved at the protein-protein interface of 2MRO^{+14}. The orange and black dots indicate the distances between the β-carbons of two residues on the protomers less than 8 Å and 4 Å, respectively. (b) Snapshots magnifying the hot spots at the interface of the protein dimer.

evaporation (Figure 4.6 (b)). To find the droplet evaporation rate, two sets of six runs of droplets composed initially of 2000 H_2O molecules were performed at 370 K and 390 K. In three of each set of the runs the complex was carrying zero charge and in the other three, it carried charge +14 *e*. At 390 K, the neutral droplet reached a size of 138−213 H_2O molecules in 5 ns, while the droplet with charge +14 *e* reached a size of 115−120 H_2O molecules in 3.5 ns. At 370 K, the neutral droplet reached a size of 187−207 H_2O molecules in 9.0 ns, while the droplet with charge +14 *e* reached a size of 237−307 H_2O molecules in 5.5 ns.

Typical snapshots of the complex dissociation followed by fission of the droplet into two sub-droplets, each of which carries a single protomer, are shown in Figure 4.7. It was found that the complex separation takes places in three distinct stages: (a) The protomers re-orient and form a new protein-protein interface (Figure 4.7 (a)), (b) the complex dissociates within the droplet (Figure 4.7 (b)), and (c) the droplet fragments (Figure 4.7 (c)).

Statistical analysis of the droplet fragmentation times over the 14 simulation outcomes was performed with the assumption that the distribution of the times ($P(t)$) follows the exponential expression

$$P(t) = \frac{1}{\tau}e^{-t/\tau} \tag{4.12}$$

where t is time and $1/\tau$ is the complex dissociation rate. By using the maximum likelihood estimate method [81], it was found that $\tau = 13$ ns. In this analysis, simple averaging of the dissociation times does not take into account the fragmentation events that do not occur beyond the cut-off time. The complexes that remained bound shift the average fragmentation rate to a value that is larger than the evaporation rate. However, the distribution of the fragmentation times suggests that there is a substantial number of complexes that they would dissociate under the particular droplet conditions (*i.e.*, elevated charge state and elevated temperature). The dissociation rate estimate can be used to find the change in the detected populations of the bound and unbound states of the complex in the gas phase relative to those in bulk solution.

To obtain an insight into protein complex dissociation and preservation, the nature of the new protein-protein interface after the reorganization of the protomers was analyzed. In the sample of the 14 realizations starting from the protein complex with the hydrophobic interface, two possible outcomes were identified within the cut-off time: (a) the two protomers reorient

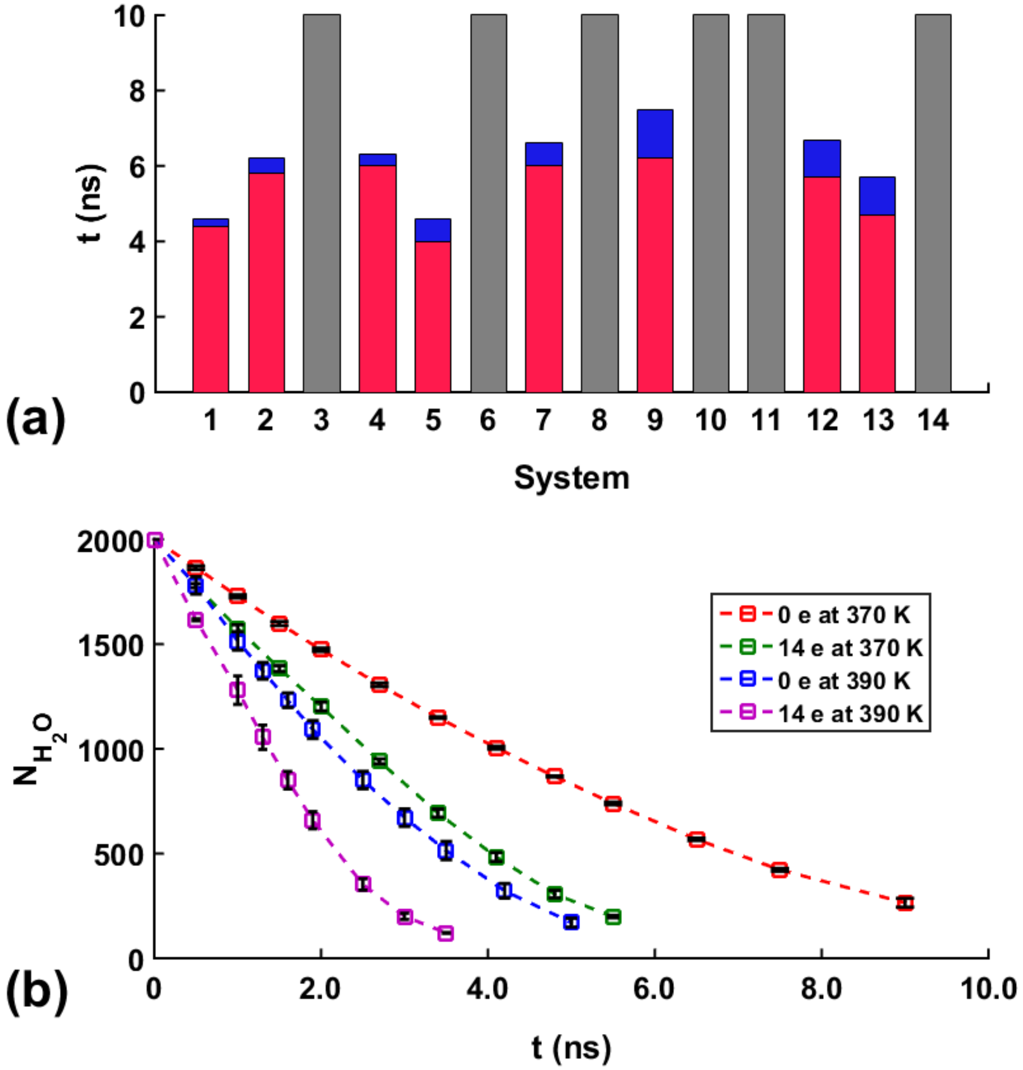


Figure 4.6: (a) Bar graph showing the time of the complex dissociation (red) and the droplet fission (blue) of a charged +14 *e* 2MRO complex in a droplet of 2000 H_2O molecules. The gray bars indicate that no complex dissociation was observed within 10 ns. (b) Number of H_2O molecules (N_{H_2O}) remaining in the droplet as a function of time (t) at 370 and 390 K. The legends show the temperature and the charge state of the 2MRO in the various evaporation runs.

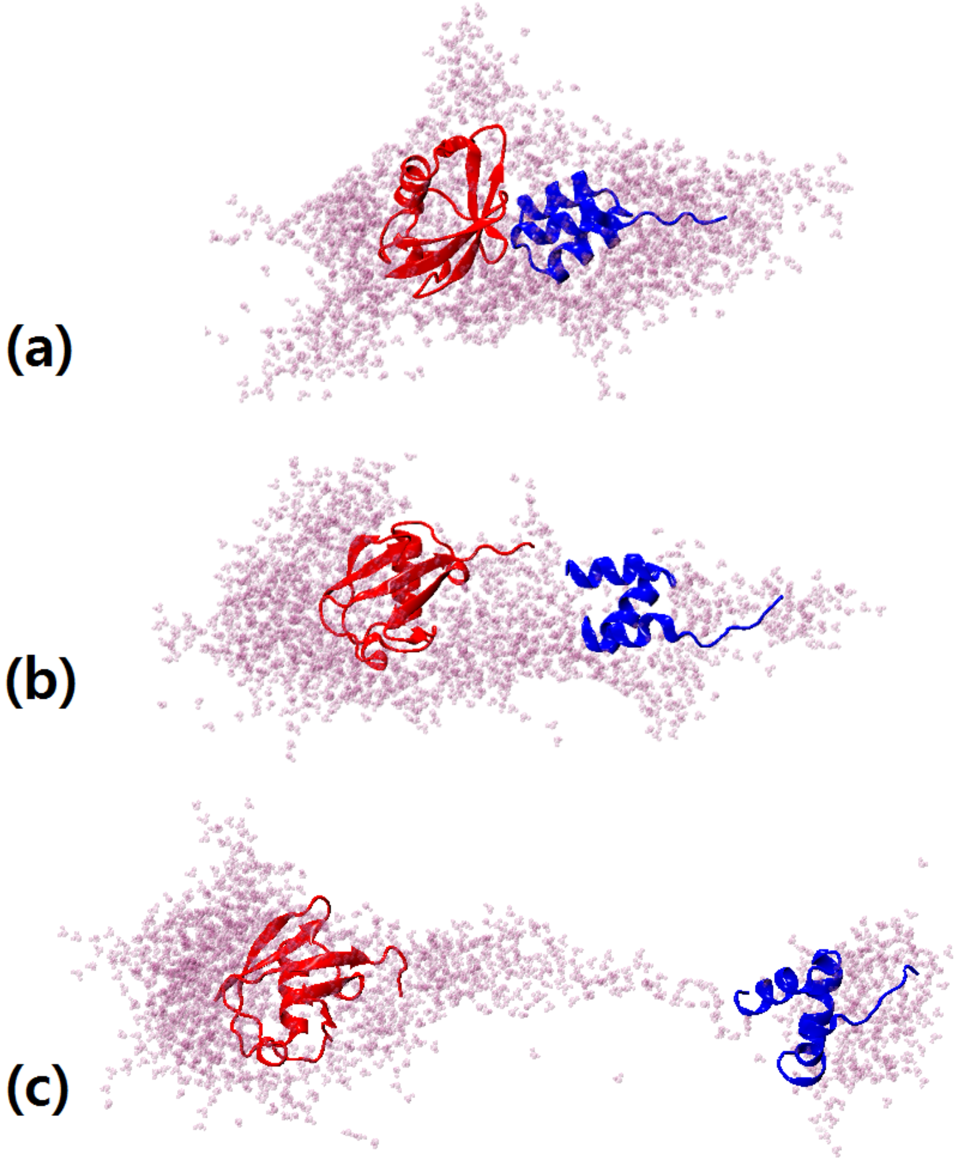


Figure 4.7: Three stages of the fragmentation mechanism of the 2MRO with charge +14 *e* in an aqueous droplet: (a) protomer reorientation, (b) complex dissociation, and (c) droplet fission. Ubiquitin is colored in red, and the ubiquitin-binding domain is blue. The red dots indicate the water molecules.

to each other and form a new interface that does not give rise to dissociation within 10 ns, and (b) the protomers reorient but form a new interface that leads to dissociation within 10 ns. The third possible outcome was found in the three additional simulations of the same droplet system at low temperature and with no net charge: (c) the 2MRO maintains its initial conformation and interface throughout the run. Figure 4.8 shows the time evolution of the root-mean-square displacement (RMSD) of the protein complex that signifies the three outcomes. The purple arrows in each snapshot are the displacement vectors that start from the centre of mass (COM) of each protomer and point to the centre of its hydrophobic domain. The purple dashed lines connect the COM's of the two protomers. Figure 4.8 (a) displays the RMSD when the protein complex has zero net charge. In this system, the RMSD converges to and then fluctuates around a certain value (approximately 4 Å) after equilibration, implying no critical changes in its overall conformation. In contrast, the RMSDs of the complex with charge +14 *e* in Figure 4.8 (b) and (c) indicate that its native structure changes considerably in two different manners. In Figure 4.8 (b), the RMSD increases rapidly, which shows with further analysis presented in the following discussion, that the complex dissociates. In Figure 4.8 (c), the RMSD increases in a step-wise fashion, for which further analysis shows the formation of a new stable protein-protein interface that does not lead to complex dissociation. The key point in the discussion is that in the droplet environment the inter-protein interaction may be modified by forming new protein-protein interfaces. In what follows, the mechanisms of the complex stabilization and dissociation will be explained.

Interestingly, in the fragmentation event of the protein complex, the area of the interface, which is another key factor that determines the lifetime of the interface, decreases during the protomer reorientation. As illustrated in Figure 4.9, the interfacial area decreases slowly during protomer reorientation, and then it drastically decreases to zero due to complex dissociation and droplet fission. The observed mechanism suggests that a combination of two order parameters which is the angle between the arrows that monitor the reorientation of the interface (Figure 4.8) and the interface area may be used along which a free energy surface may be computed.

The likelihood of the complex dissociation depends largely on the interactions found in the new protein-protein interface. In order to understand what residue pairs are involved in the formation of the interface, the contact map of every residue in each system after reorien-

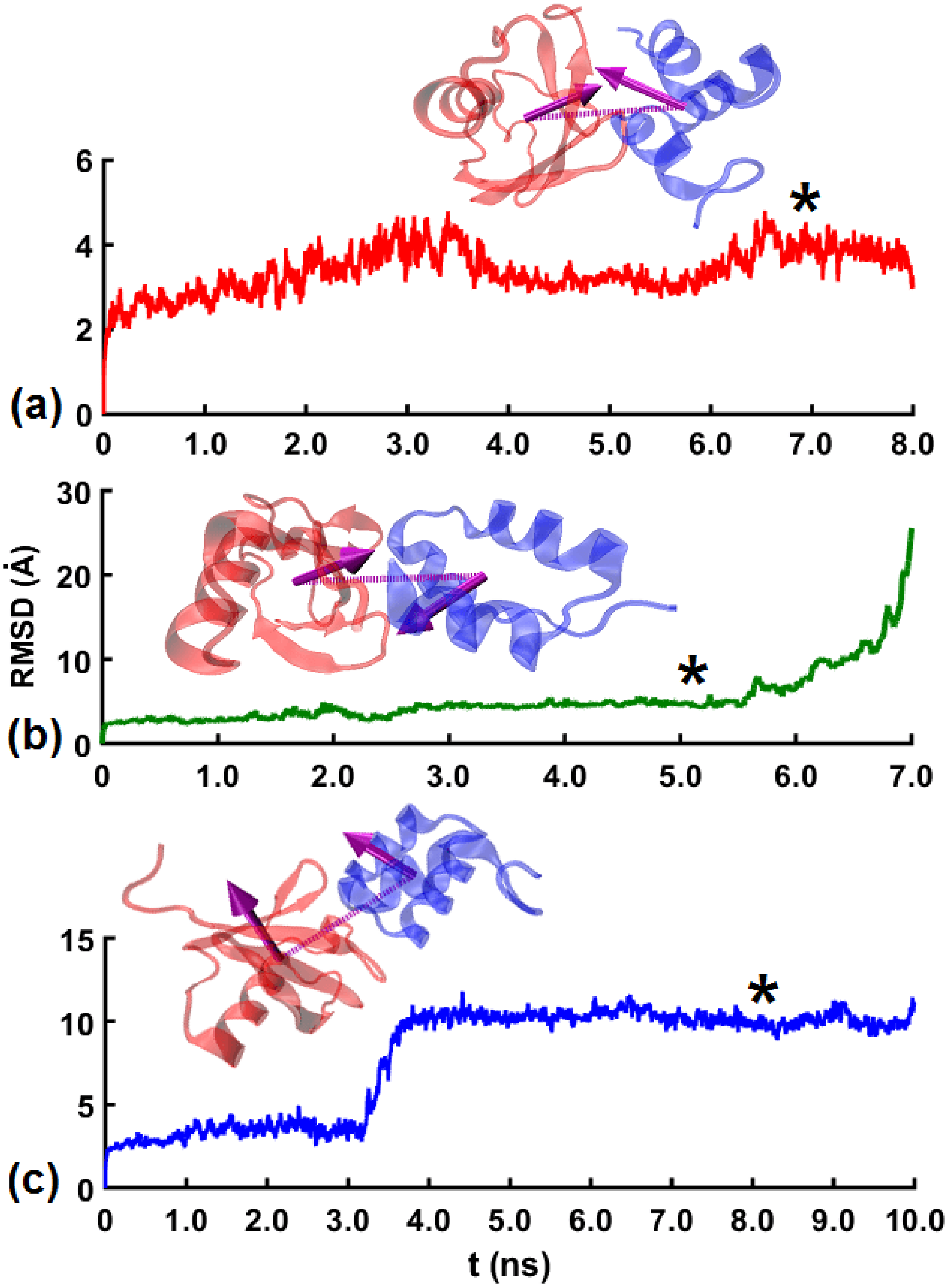


Figure 4.8: Root-mean-square displacement (RMSD) as a function of time (*t*) for three different cases of the transient protein complex in an aqueous droplet: (a) the complex is electrically neutral and (b-c) with a protonation state of +14 *e*. The protein complex in (b) dissociates, whereas the protein complexes in (a) and (c) do not. The snapshot of the protein complex at *t* marked with an asterisk for each case is also shown.

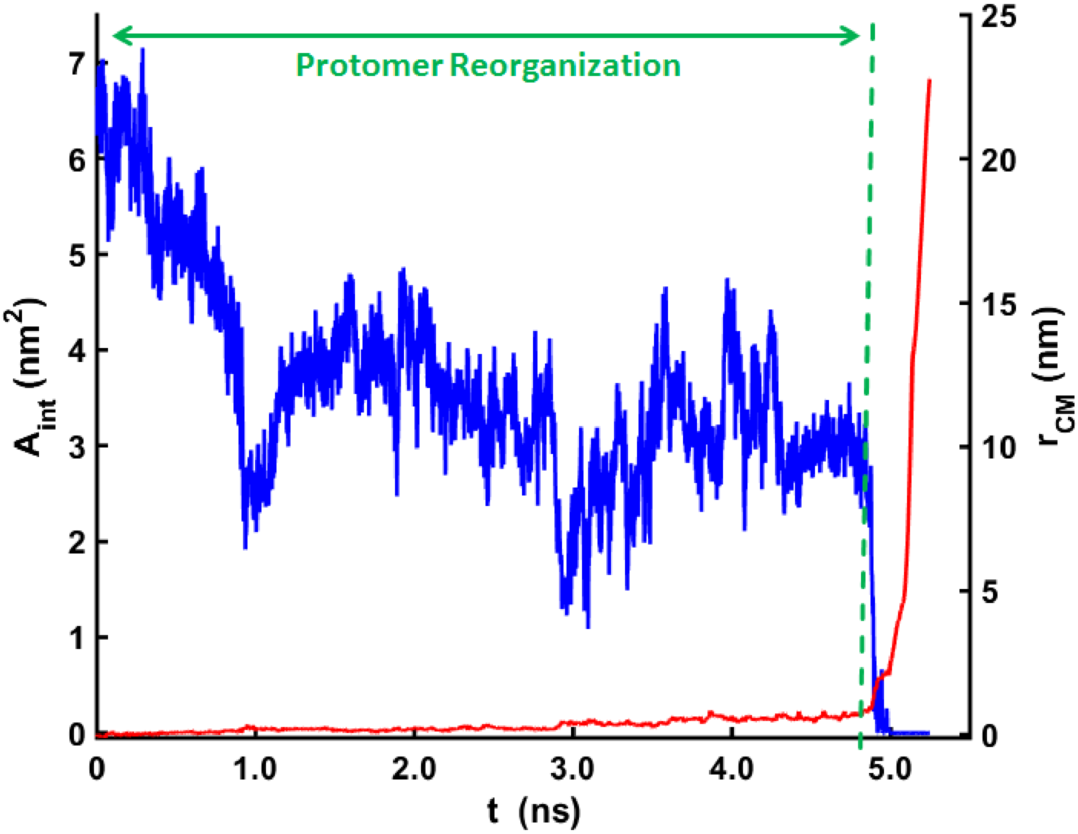


Figure 4.9: Interfacial area (A_{int}, blue) and separation distance (r_{CM}, red) as a function of time (t) for a dissociating protein complex. During protomer reorientation, there is a decreasing trend of the interfacial area and an increasing trend of the separation distance (before the green bar). After complex dissociation, the interface area is decreasing sharply to zero, while the separation distance abruptly increases.

tation was drawn and analyzed. For the analysis, each residue is classified into two groups: hydrophobic and polar/charged residues. The contact maps reveal that the occupancy of polar/charged residues at the interface is higher for the complexes that undergo no dissociation after protomer reorientation. The frequencies of hydrophobic and polar/charged residues at the interface are 48.5 % and 51.5 %, respectively, for the dissociating protein complexes, while they are 33.5 % and 66.6 %, correspondingly, for the stable complexes. Here, it is clearly seen that in both cases, polar/charged residues are more frequently found at the interface than hydrophobic groups. This is strongly expected since the number of surface charged/polar amino acids were initially widely spread over the complex surface other than the hydrophobic patches. However, when the number of residues involved at the interface is very small, the frequency is overestimated and thus the type and number of interactions become less significant. The intermediate complex can be stabilized in some conformations. This is, after all, mainly due to the contributions from electrostatic interactions. The Coulombic attraction is known to prolong the lifetime of the encounter complex and reduces the surface area to be searched for binding [82]. This also indicates that the binding of the protomers does not guarantee the preservation of the initial protein-protein interface and thus the overall conformation. Based on the simulation data, it was found that there exist two different main surfaces on each constituent protomer forming an interface (Figure 4.10). It is worth noting that the interface that results in complex dissociation shares a common region with the hydrophobic patch to great extent (red). The green region harbours more polar/charged residues and stabilizes the intermediate complex. The other region coloured in gray is not searched by the partner in all the simulations. In summary, we found that the stabilization of new protein-protein interfaces after protomer reorientation is mainly due to the contributions from electrostatic interactions. The contact map for the two distinct interfaces found in 2MRO is presented in 4.11.

In bulk solution, 2MRO in the charge state of +14 e did not dissociate within 34 ns at T = 390 K. The time evolution of RMSD is shown in Figure 4.12. The RMSD indicates that the protein-protein interface may be destroyed but it is reformed without complex dissociation. The data suggest that the complex dissociation occurs faster in the nanodrop relative to the bulk solution. In this particular charge state, the increased complex dissociation rate is attributed to the limitation of the maximum charge that a droplet can sustain. In other words, it appears as a

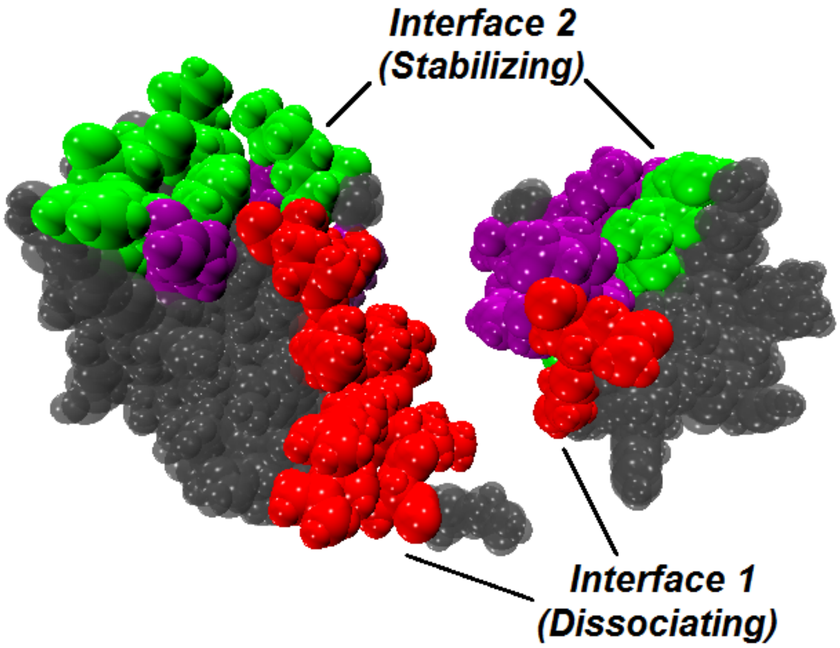


Figure 4.10: Interfaces found in the 2MRO complex in an aqueous droplet. The water molecules are not shown. The interface colored in red leads to complex dissociation after protomer reorientation, whereas that in green causes stabilization of the intermediate complex. The purple colour indicates the region shared by both red and green patches. No interface formation is found in the gray region. The region in the various interfaces was compiled by the findings in the fourteen simulation runs.

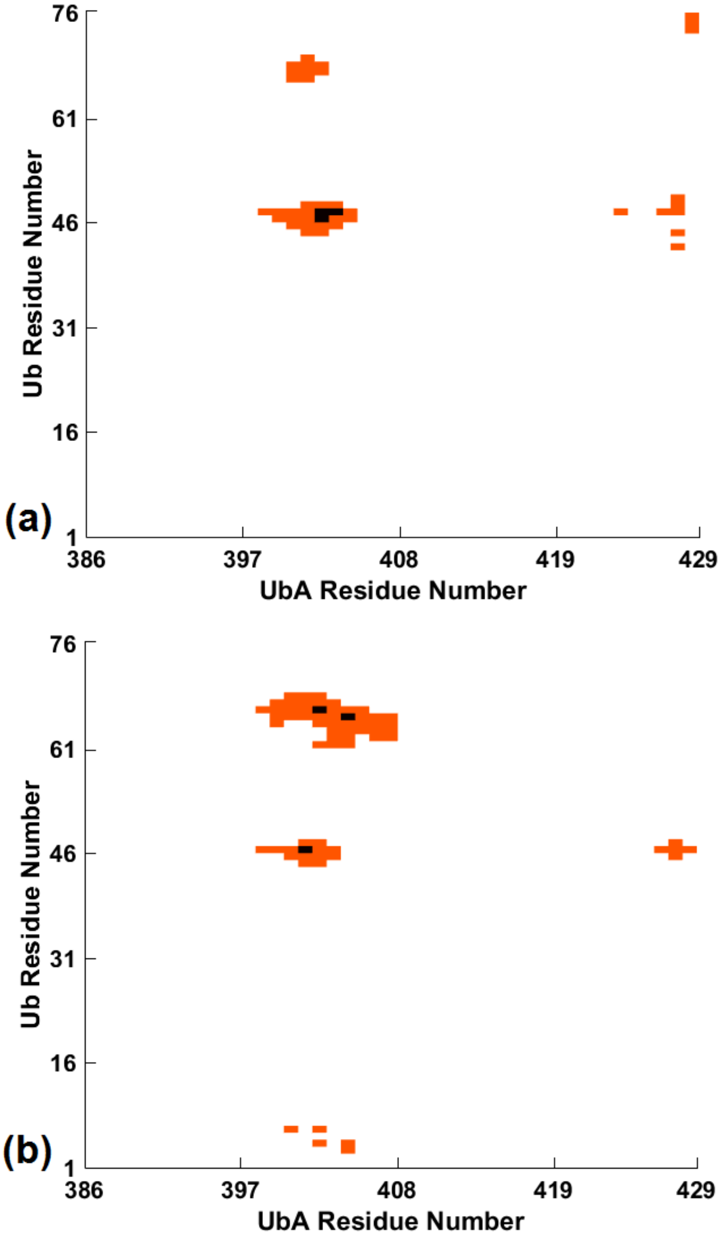


Figure 4.11: Example of the contact maps showing the type of residues (indicated by their residue numbers) that are involved in the interface of an intermediate complex that leads to (a) dissociation and (b) stabilization after protomer reorientation. The orange/black colouring scheme is the same as in Figure 4.5.

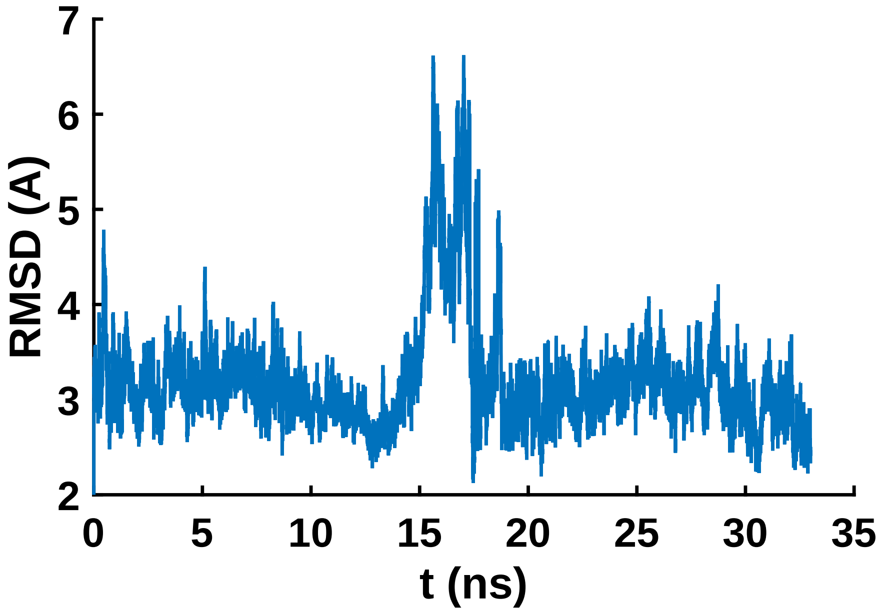


Figure 4.12: Root-mean-square displacement (RMSD) as a function of time *t* for 2MRO^{14+} with the hydrophobic protein-protein interface in bulk solution at 390 K.

form of charge-induced instability of the droplet.

We found that the complex dissociation in our simulations is accelerated by the combination of the conditions of high temperature and of high charge state of the complex in a droplet. This may not be surprising but it was not obvious before the simulations were performed. The presence of one of the conditions only is likely to prolong the bound state of the complex in a droplet. For example, as shown in Figure 4.8 (a), no complex dissociation occurs when the total charge of the complex is zero. Also no complex dissociation is detected when 2MRO has +10 *e* in the same size of the aqueous droplet. Therefore, the charge reduction of the protein complex may extend the lifetime of the bound state. At 300 K, the aqueous droplet composed of 1400 H_2O molecules and 2MRO^{14+} does not shift the protein complex to the unbound state within 15 ns. Here, both the stabilizing and destabilizing interfaces were tested. No substantial conformational change takes place for the stabilizing interface, whereas protomer reorientation occurs for the destabilizing one. The RMSD plots are presented in Figure 4.8.

The elucidation of the mechanism of the complex dissociation may propose at least two different approaches to be used to preserve the non-covalent interactions: (i) connecting the two protomers via, for instance, cross-linking [1, 83] or (ii) relieving the charge-induced instability and so preventing droplet fission by adding counter-ions. These counter-ions must be soluble in water and solvated in the interior of the droplet. These ions can bind to the protein complex to reduce the net charge of the protein complex and so the droplet.

4.3.3 Effect of Charge Distribution on the Complex Stability

According to the guiding principles presented in Section 4.2.1 that determine the protonation state of the protein complex, it was found that the charge distribution of +11 *e* in Ub and +3 *e* in UbA is the most prevalent one in the ensemble of the droplet systems under the specified conditions. However, there should be small populations of different charge distributions in the ensemble because the same acidic residues may have slightly different pK*a* values depending on their chemical environment. To gain an insight as to how charge distribution may affect the dissociation mechanism of the protein complex, we performed MD simulations of the protein complex with different charge states and initial protein-protein interfaces: (+12 *e* in Ub and +2 *e* in UbA *vs*. +10 *e* in Ub and +4 *e* in UbA) and (hydrophobic interface *vs*. destabilizing interface).

When the 2MRO complex ion has the hydrophobic interface between the protomers, no protomer reorientation is detected and so the initial interface is maintained during the evolution of the droplet system (namely, for longer than 10 ns). In the simulation of the +12 *e*/+2 *e* charge distribution, the enhanced stabilization of the interface is attributed to the presence of an additional salt bridge as indicated by the black square in Figure 4.13 (a). In this charge distribution compared to the +11 *e*/+3 *e* charge distribution, the carboxylic group of the terminal residue (serine-429, coloured in green in the same figure) of the UbA domain is deprotonated bearing a negative charge, while aspartate-58 (coloured in purple in the same figure) of ubiquitin is protonated neutralizing its negative charge. Here serine-429 is located at the rim of the hydrophobic interface and interacts with two positively charged basic residues (arginine-42 and arginine-72, coloured in yellow in the black box) positioned in proximity to serine-429.

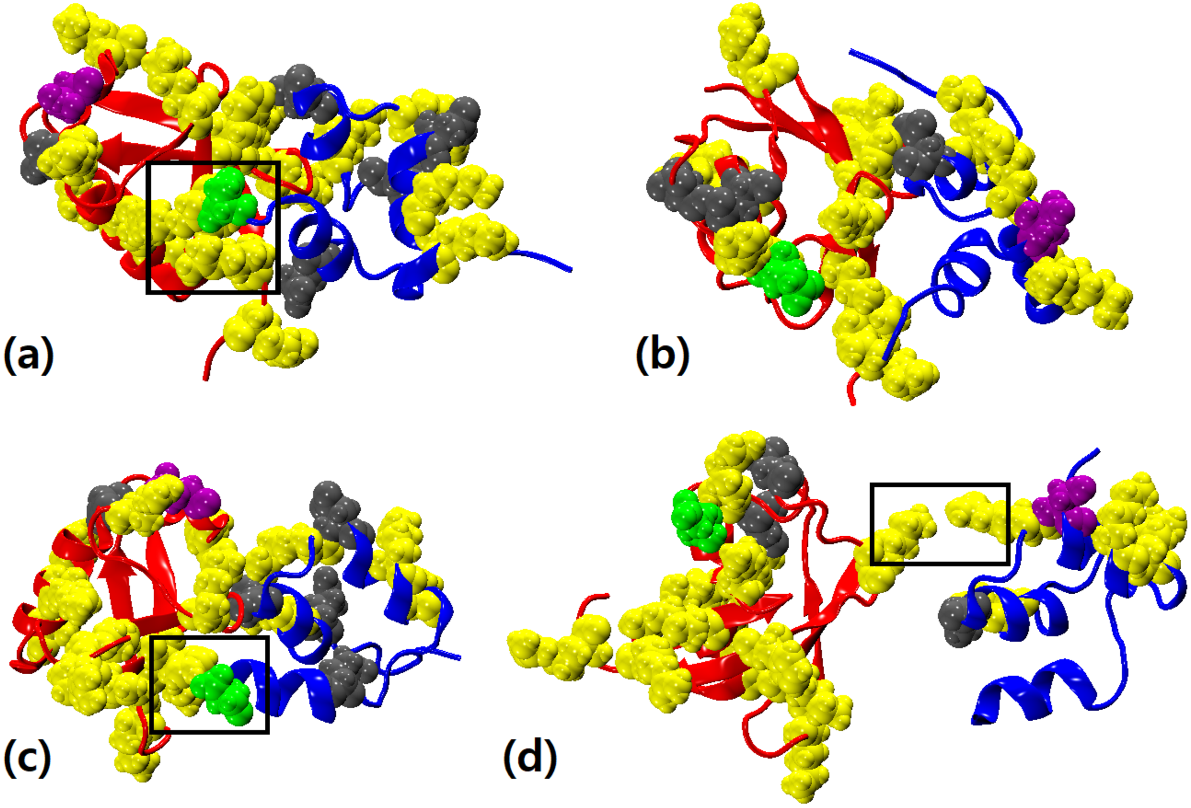


Figure 4.13: Snapshots of the 2MRO^{14+} ion with different charge distributions and initial protein-protein interfaces: (a) +12 *e*/+2 *e* and hydrophobic interface, (b) +10 *e*/+4 *e* and hydrophobic interface, (c) +12 *e*/+2 *e* and destabilizing interface, and (d) +10 *e*/+4 *e* and destabilizing interface. Ubiquitin and the UbA domain are colored in red and blue, respectively. Positively charged basic residues and negatively charged acidic residues are colored in yellow and gray, respectively. The purple and green colors indicate the residues on ubiquitin and the UbA domain, respectively, which changes its protonation state in comparison to the 2MRO^{14+} ion with +11 *e*/+2 *e* charge distribution. The residues grouped by a black square are the ones that determine the stability of the protein complex (see the main text).

The salt bridges formed among these three residues contribute significantly to the "locking" of the protein complex conformation, preventing complex dissociation during the lifetime of the droplet. It is important to notice that this outcome has not been observed in the 2MRO with +13 *e*/+3 *e* charge distribution. On the contrary, aspartate-58 is far away from the protein-protein interface and thus it provides no or little contributions to the stabilization of the interface. In the +10 *e*/+4 *e* charge distribution relative to the +11 *e*/+3 *e* charge distribution, aspartate-406 (coloured in green in Figure 4.13 (b)) of the UbA domain is deprotonated possessing a negative charge, whereas glutamate-51 (coloured in purple in the same figure) of ubiquitin is protonated possessing a zero charge. Since these two resides are located far from the hydrophobic interface, its direct role in the stability of the interface is not expected. In other words, the major interactions involved at the protein-protein interface in this charge distribution is highly similar to those found in the +11 *e*/+3 *e* charge distribution. The only difference is then the amount of net charge carried by each protomer.

When the 2MRO complex ion has the destabilizing interface between the protomers, protomer reorientation takes place but it does not secure complex stabilization. In the simulation of the +12 *e*/+2 *e* charge distribution, protomer reorientation results in a new protein-protein interface stabilized by the addition of a salt bridge between histidine-68 (coloured in yellow in the black box in Figure 4.13 (c)) of ubiquitin and the carboxylic group of serine-429 (coloured in green in the same figure) of the UbA domain. The protein complex maintains the conformation for the rest of the droplet's lifetime. This bears a resemblance to the scenario observed in the 2MRO complex ion with the stable PPI after protomer reorientation (*i.e.*, Figure 4.8 (c)). In the simulation of the +10 *e*/+4 *e* charge distribution, in contrast, it was found that complex dissociation readily takes place within 300 ps, followed by droplet fission after 700 ps. This is quite expected because the interface should be of high similarity to the destabilizing interface found in the 2MRO complex with the +11 *e*/+3 *e* charge distribution; as aforementioned, aspartate-406 (coloured in green in Figure 4.13 (d)) and glutamate-51 (coloured in purple in the same figure) are not directly involved in the formation of the interface. The simulation shows that complex dissociation is initiated when the two positively charged basic residues in the black box in the same figure, lysine-63 (of ubiquitin) and arginine-405 (of the UbA domain), approach and repel each other.

Therefore, both the charge distribution (that is, the type of interactions involved at the interface and the amount of charge carried by each protomer) and the relative orientation of the protomers may play a critical role in the stability of a protein complex ion in a droplet.

4.3.4 Dynamics and Mechanism of $2PEA^{14+}$ Dissociation

The stability of the 2PEA complex ion with net charge +14 *e* (equally distributed to the two ubiquitin molecules) in an aqueous droplet was examined in order to examine whether some of the features of protein dissociation that we found in 2MRO also hold for other weak transient complexes. The protein complex is solvated in a droplet composed initially of 833 H_2O molecules. After equilibration, $\approx 650-750$ H_2O molecules remain in the main droplet. Since the 2PEA is equivalent in volume to $\approx 750-770$ H_2O molecules, the Rayleigh limit of the droplet system is approximately $+9-10$ *e*. Therefore, the spherical water droplet-$2PEA^{14+}$ system is substantially above the Rayleigh limit.

The group of 10 MD simulations for the system reveal the same three-step pathway in its dissociation and droplet fragmentation. In other words, complete separation of the ubiquitin protomers is achieved via protomer reorientation, complex dissociation, and droplet fission.

The destabilization process of the protein-protein interface in the 2PEA complex is illustrated in Figure 4.14 (a). In the plot, the RMSD increases gradually during protomer reorientation and swiftly at the moment of complex dissociation. Here the pattern of the RMSD plot of the protein dimer is the same as that in found Figure 4.8 (b) where the 2MRO complex ion dissociates after protomer reorientation. The conformations of the protein complex at different times are shown in the snapshots embedded in the graph. The formation of a new protein-protein interface during the evolution of the droplet system is verified by the motions of the purple arrows, each of which points from the COM of the ubiquitin molecule to the centre of its hydrophobic patch.

Figure 4.14 (b) shows all the charged amino acids located at the surface of the protein complex. Every deprotonated acidic residue bearing $-e$ is coloured in yellow, while every protonated basic residue having $+e$ is coloured in gray. Since the positively charged basic residues are concentrated around the hydrophobic interface between the two ubiquitin molecules, the destabilization of the interface is assisted by electrostatic repulsion between their positive charges. After protomer reorientation, the protein molecules are connected by the unstable interface involving several protonated basic residues (*e.g.*, lysine-48 and arginine-72, coloured in light blue in the same figure), which accelerate the separation of the protomers by repelling each other. Other positively charged basic residues commonly found in the destabilizing interface include arginine-42, arginine-54, and lysine-11 of each ubiquitin molecule.

Complete complex separation takes place within 0.2−5.3 ns in the 10 simulation runs. Table 4.1 summarizes the times required for complex dissociation and droplet fission. A non-equilibrium run of the droplet evaporation was carried out to find an upper time limit for terminating the equilibrium runs in the cavity, which indicates that the droplet system reaches the size of approximately 340 H_2O molecules within 6 ns. Therefore, the protein complex would likely dissociate before its desolvation and transfer to the vapour phase. Here, it is noticeable that the rate of complex dissociation, if occurs, is higher for $2PEA^{14+}$ than $2MRO^{14+}$. This is because the ratio of the net charge to the critical limit of the droplet is much larger for the system with $2PEA^{14+}$ (1.08-1.27) than the system with $2MRO^{14+}$ (1.40−1.55). Moreover, the time lapse between complex dissociation and droplet fission is much shorter for the system with $2PEA^{14+}$ ($\leq$ 50 ps) than the system with $2MRO^{14+}$ (0.2−1.3 ns). This is attributable to a fewer layers of water molecules solvating the diubiquitin complex ion.

Based on the multistep mechanism of complex dissociation that is consistently observed in 2MRO and 2PEA, at least two free energy barriers are envisioned along certain order parameters. A water droplet that holds a noncovalently bound complex must overcome the first energy barrier required for “twisting” the protein complex to have a dissociating protein-protein interface. The buried surface area (BSA) of a protein complex is a hydrophobic surface removed from contact with solvent and buried in the interface. BSA is a descriptor related to the binding affinity of a PPI, whose magnitude is estimated to be approximately 0.025 kcal/mol per 1 $\mathring{A}^2$ of the hydrophobic surface [84]. It was calculated that the 2MRO complex has $\approx$ 624 $\mathring{A}^2$ which is in the range of a typical transient protein complex [41]. Based on the estimation, the binding free energy at the hydrophobic interface is estimated to be $\approx$ 15.6 kcal/mol. Then the second energy barrier appears, which arises from asymmetric fission of the droplet; we found that the smaller progeny droplet draws 12−25 % of water from the parent droplet. From the manner in

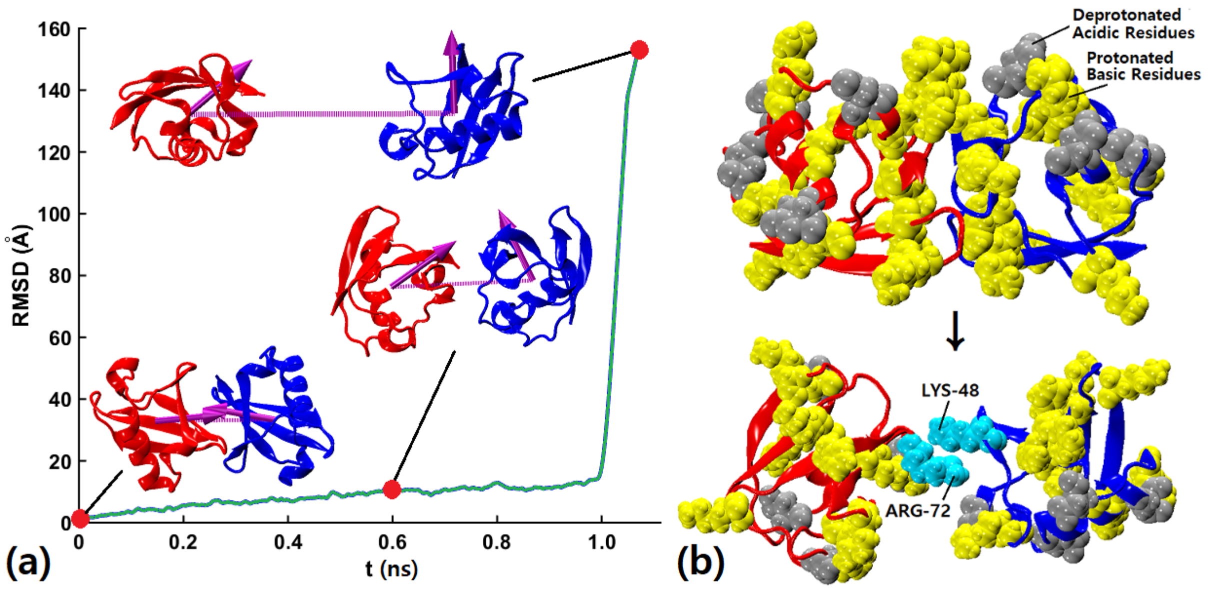


Figure 4.14: (a) Root-mean-square displacement (RMSD) of 2PEA^{14+} as a function of time (t) with the snapshots corresponding to the protein complex at the moments marked with red circles on the curve. The two ubiquitin molecules are colored in red and blue. The purple arrow is the displacement vector connecting the center of mass (COM) of each ubiquitin to the center of its hydrophobic patch. The dashed purple line connects the COM's of the two ubiquitin molecules. (b) Typical snapshots that show protomer reorientation of the 2PEA^{14+} complex ion. The deprotonated acidic residues and the protonated basic residues are colored in gray and yellow, respectively. The residues highlighted in light blue indicate the locations of lysine-48 and arginine-72 at the destabilizing interface of the 2PEA^{14+} ion right before complex dissociation.

Table 4.1: Times required for complex dissociation (t_d) and droplet fission (t_f) following the complex dissociation. The systems are composed of droplets with approximately 650-750 H_2O molecules and a 2PEA ($\approx$ 750-770 H_2O molecules equivalent in volume). CT and CE represent constant-temperature runs and constant-energy runs, respectively.

System	t_d (ns)	t_f (ns)	System	t_d (ns)	t_f (ns)
CT1	0.99	1.00	CE1	0.24	0.29
CT2	0.37	0.40	CE2	0.14	0.18
CT3	0.49	0.51	CE3	0.32	0.36
CT4	0.95	0.97	CE4	1.40	1.43
CT5	0.68	0.72	CE5	5.29	5.31

which the droplet fission takes place (Figure 4.7 (c)) the bottle-necked droplet shapes are best captured by the "transfer" order parameter described by Consta *et al.* [85, 86] for release of ions from droplets.

4.3.5 Effect of "Star" Morphologies of Droplets on Protein Complexes

Even though the "star"-shaped droplets have not been observed experimentally as yet, we think that they may play a role in determining the charge state and the stability of a protein complex in an evaporating droplet.

In relation to ESI experiments, it is very likely that "spiky" structures appear in the course of droplet evolution from bulk solution to the gaseous state. The origin of the charge states of globular proteins and their complexes has been a fundamental question in ESI-MS and as that, it has been discussed in the literature over several decades [87–93]. Of course, the charge state of a protein can be manipulated by changing the experimental conditions such as the surface tension of the droplets by using different solvents and co-solvents, or various buffers. A collection of the most abundant charge state and the maximum charge state of a variety of proteins measured in the positive and a few in the negative ion mode of ESI is presented in Table 1 [92] of the article of [92]. These measurement are made in the Heck laboratory by nano-flow ESI from 50 mM aqueous ammonium acetate at neutral pH. A model proposed by de la Mora [89] has often be used to explain the final charge state of a protein that is transferred into the gaseous state by the ESI process. The model assumes that in the latest stage of droplet evaporation, when a droplet comprises the compact macroion and a layer of water surrounding the macroion, the droplet will hold as many charges as predicted by the Rayleigh model. The model states that further solvent evaporation will collapse the ions on the compact protein, therefore, the charge on a globular protein will be equal to the charge predicted by the Rayleigh model for a liquid droplet of the size of the protein. Using the Rayleigh model ($X = 1$ in Equation 1.6) de la Mora provided an empirical relation $Q_R = 0.078M^{0.5}$ that relates the protein charge Q_R with its molar mass M. Figure 2 in the article of [92] shows a good agreement between the de la Mora model and the average charge state of the proteins detected in the positive ion mode. This model assumes that the mass density of the globular protein is the same as that of water and it also implicitly assumes that the last layer of water surrounding the macroion is uniformly spread on the surface of the protein. The second assumption may not be always true due to the inhomogeneity of a protein surface. Transfer of the charge on the macroion when the last layer of solvent surrounds the macroion may be one of the mechanisms that a globular protein acquires its charge. The deprotonation of a protein is also a possible mechanism [42].

We found in the literature that several proteins have been reported with an average charge state higher than that predicted by de la Mora's model. Examples of such proteins are the para-hydroxybenzoate hydroxylase in Ref. [93], 2PEA in Ref. [73], and urease $\alpha_{18}\beta_{18}$ and urease $\alpha_{24}\beta_{24}$ in Ref. [92]. Moreover, the maximum charge state of a protein (which is different from the average charge state) is often above the Rayleigh limit of the corresponding spherical droplet. In Section 4.3.4, we have examined a diubiquitin complex, which has been detected at the charge state of +14 e by using cryo-IM-MS [73]. The charged 2PEA was formed under native ESI conditions. We applied the empirical relation $Q_R = 0.078M^{0.5}$ in the ubiquitin dimer and found that it predicts that the final charge state of the dimer would be likely +10 e. The comparison between the experimental findings and the empirical observation implies that, regardless of the manner in which the protein obtained the charge state, the entire droplet system should have held at some point during the course of solvent evaporation more charges than the critical value predicted by the Rayleigh criterion (for the corresponding spherical droplet). The only way this could have happened is through a deformed "star" shaped droplet. Indeed, we found as shown in Figure 4.15 that the droplet morphology is characterized by conical protrusions of water.

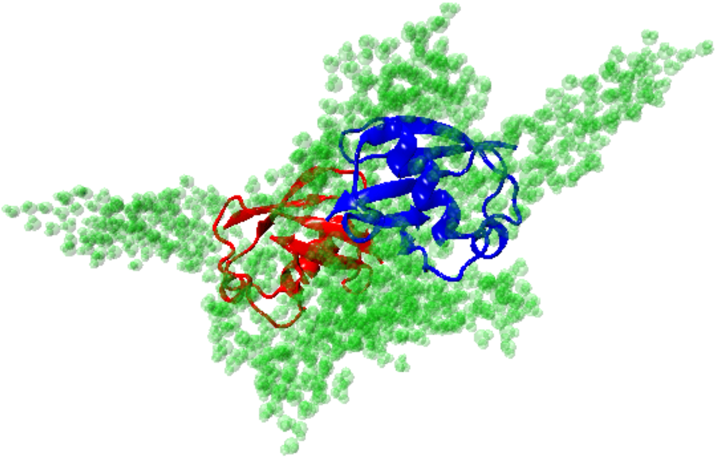


Figure 4.15: Typical snapshot showing the star formation of a charged aqueous droplet which contains 2PEA with +14 *e* and 833 H_2O molecules. The ubiquitin molecules are coloured in red and blue, and the water molecules are coloured in transparent green.

The finding lead us to suggest the following two scenarios for the charging of 2PEA. (a) The 2PEA had acquired the charge state of +14 *e* at an earlier state [42] of the droplet evolution before the final layer of solvent remained on the surface of 2PEA. In this earlier state a larger droplet may had contained the charged protein and co-ions. As the droplet shrinks co-ions are released and the only ion left in the droplet may be the macroion. We argue that, the droplets may be found in stable "spiky" states above the Rayleigh limit of the corresponding spherical droplet without shedding away hydronium ions. We think that this occurs because the formation of spikes provides certain pathways for ion release. The ions can be released only via these pathways and not from the troughs. Considering that the solvent conical protrusion may not extend out from the charged amino acids and also that their location changes via birth and death processes of spikes (as shown in the next paragraph), their presence may inhibit the proton release from the protein. (b) If the droplet has a charged 2PEA (in a charge state below +14 *e*) and simple ions, it is still very likely for the droplet to have a transient "spiky" shape. At this state, the release of simple ions is delayed [94] relative to a droplet with only simple ions. If the co-ions are ammonium ions, proton transfer reactions from the co-ions to the protein may occur. Based on our analysis, we propose that the formation of "spiky"-shaped droplets is part of the mechanism that explains the charge state of a protein.

Figure 4.7 shows the star formation of the aqueous droplet charged with the 2MRO complex ion (+14 *e*), where three or four rays are developed around the macromolecule. It was found that a protein-vapour interface may also form as a consequence of this charge-induced instability during or before the transfer of the macromolecule from the droplet to the gas phase. These spikes may not be very static both in time and location due mostly to high thermal energy and protomer reorientation. It was observed that different numbers of spikes are developed on the droplet surface during the evolution of the droplet, as shown in Figure 4.16. The lifespan and the location of those protrusions can vary. High temperature enhances surface fluctuations of the droplet, generating short-living spikes as indicated by a dashed black circle in Figure 4.16 (c). They may appear in different locations on the droplet surface. On the contrary, the two spikes pointing outward from each protomer (indicated by dashed squares in the same figure) are long-living and rather fixed in location.

Direct visual inspection shows that some residues located at the protein surface may be exposed bare of solvent into the vapour phase when the system is in the regime of charge-induced instability. Intriguing enough, this is not observed when the protein complex has no net charge and thus is well solvated, due to the absence of spiky solvent protrusions around the macromolecule. As water molecules are merely dragged to the formation of a star-like morphology, some of the amino acids at the surface of the protein complex are exposed directly to the vapour phase, particularly when those residues are hydrophobic and at the trough between two adjacent thorns of the star. These residues can be identified by plotting the number of water molecules in the first solvation shell ($N_{\mathrm{H_2O}}$) and the solvent accessible surface area (SASA) for each amino acid comprising the protein complex. It is clearly seen that there is a positive, linear correlation between those two quantities; as the SASA increases, the degree of hydration tends to rise. The red crosses scattered within the blue box in Figure 4.17 (a) correspond to the amino acids that are not well solvated. For example, those sited in the hydrophobic core of each protomer and at the interface are located in this region as water molecules have no or limited access (*i.e.*, very small SASA). However, the red crosses enclosed by the red circle show the existence of few water molecules around the corresponding residues despite higher chances of solvation (*i.e.*, relatively large SASA). This enhancement in dewetting is observed for alanine, valine, phenylalanine, serine, and leucine located at the surface of the protein complex as shown in Figure 4.17 (b). In terms of a hydrophobicity scale of an amino acid, phenylalanine and leucine are

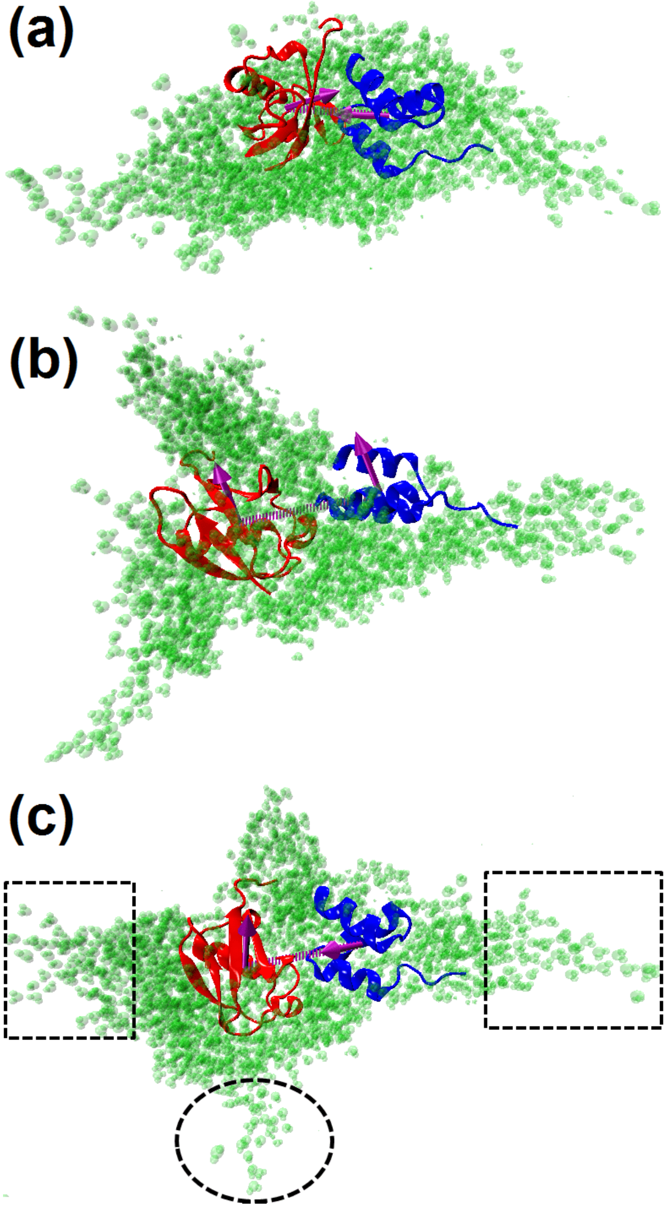


Figure 4.16: Different star morphologies of the aqueous droplet that contains 2MRO with +14 *e* and ≈ 1400 H_2O molecules. The same colouring scheme for the protein complex and the arrows was used as in Figure 4.15, and the water molecules are coloured in transparent green. The transient spikes and the long-living ones are indicated by a dashed black circle and a dashed black square, respectively.

usually at the apex, followed by alanine with a moderately high value, whereas serine has a polar neutral side chain with mild hydrophilicity. Serine is often classified as a water-friendly amino acid, but in this case it is enveloped by the other four hydrophobic amino acids and thus takes part in the bare exposure to the vapour phase. The dewetting of some surface residues may be regarded as a significant effect of charge-induced instability of a droplet, since it provides a distinct chemical environment for those and neighbouring residues due to the partial or complete loss of solvent access. The partial solvation of the complex is not expected to have a direct influence on the stability of the protein-protein interface. The restriction on the access of solvent will create a unique chemical environment for the amino acids. For example, where co-ions and counter-ions may exist is the region where the solvent is accumulated, which may affect the stability of a protein complex [95–97].

4.3.6 Estimation of Errors in K_d due to Complex Dissociation in a Droplet

In this section, the overall picture of the methodology used in this study is summarized and linked to the derivation of a significant expression that estimates the error in the measurement of K_d using ESI-MS.

Typically ESI protein complexes are found in acidic aqueous droplets (*i.e.*, in a positive ion mode) [98]. The smaller the droplet, the lower the droplet pH. Among other factors, the droplet acidity is determined by the concentrations of NH_4^+ and CH_3COO^- ions. Even in the absence of ammonium acetate, the droplet environment is still acidic because of electrolytic processes that occur in the ESI instrument. In order to study the stability of a protein complex, in principle, one may monitor the droplet evaporation in a large number of simulation runs starting from different initial conditions of the droplet-complex system, and analyze the final state of the complex. A major bottle-neck in performing those simulations is, however, that the protonation state of the protein may change during the droplet evaporation since the pH of the droplet constantly changes. The protonation state of a protein at fixed pH has been a major computational problem that is being studied over several decades in bulk solution [49–59]. In order to consider the protonation state of a protein or its complex in a shrinking droplet where the pH decreases, we devised the following method: Equilibrium MD runs of a droplet

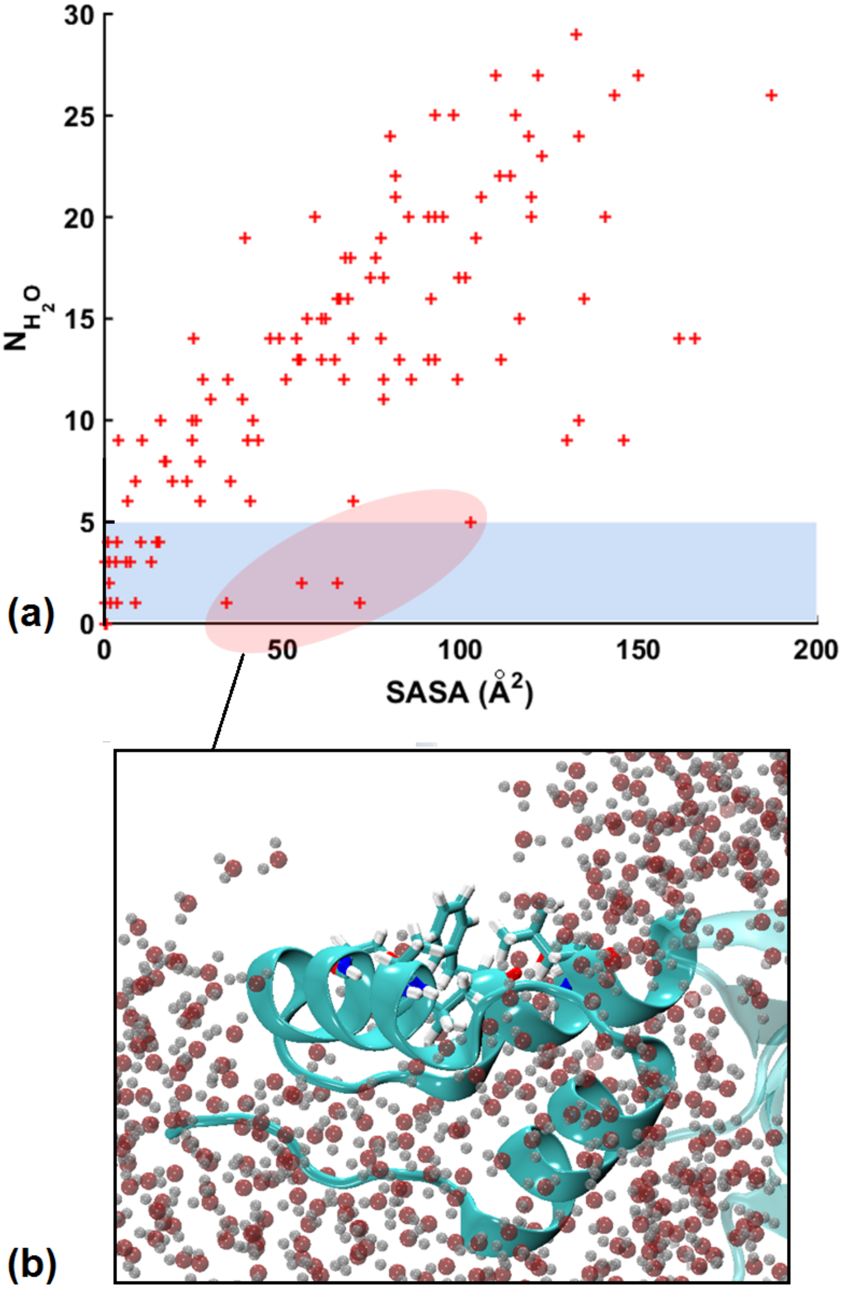


Figure 4.17: (a) The number of water molecules within 5Å (N_{H_2O} and the solvent accessible surface area (SASA) of each amino acid (marked with a red cross) in 2MRO^{14+} at a particular time frame. (b) The corresponding snapshot showing enhanced dewetting. The oxygen and hydrogen atoms in a water molecule are represented by transparent red and white spheres. The figure shows part of the surface of the protein complex where alanine, valine, phenylalanine, serine, and leucine are exposed to the vapour phase.

of specific size are performed instead of non-equilibrium runs. The equilibrium is achieved by placing the droplet in a spherical cavity. The cavity is a spherical potential that confines evaporating molecules within a volume. Water evaporation from the droplet in the vacant space of the cavity establishes an equilibrium vapour pressure. For the particular droplet size, the pH and the corresponding protonated state of the protein are determined based on the method presented in the previous sections. We can perform simulations of different droplet sizes at the corresponding pH and complex protonation state. Since the simulations are at equilibrium, the reactivity of NH_4^+ or CH_3COO^- ions do not change the pH. The droplet is first thermalized at the desired temperature by using a thermostat. One may turn off the thermostat and continue the simulations in the microcanonical ensemble (constant energy, volume, and number of molecules). The question that arises now is: how long should the MD run be? In order to determine the length of the MD run (that is, the cut-off time t_{cutoff}), one estimates the rate of droplet evaporation from a separate run. From the evaporation rate, one estimates the time required for the droplet to change its size. This is the cut-off time of the equilibrium MD runs in the cavity. We monitor whether there is protein complex dissociation and subsequent droplet fragmentation within the MD time period. If the dissociation rate is comparable or higher than the evaporation rate, it is deduced that protein dissociation may occur in the droplet and it may alter the value of the equilibrium constant of the complex. The protein dissociation times are fitted to a decaying exponential function from which we determine the complex dissociation rate. By assuming first-order kinetics, we can estimate the number of dissociated complexes by comparing the complex dissociation rate (τ) with the desolvation rate (t_{cutoff}), and thus, the error in the experimentally measured equilibrium constant due to the complex dissociation in the droplet carrier.

Suppose one performs N equilibrium MD simulations of a protein complex in a nascent droplet. Using the methodology presented in this section, one can compute the fraction of dissociating complexes during the lifetime of the droplet. The fraction should converge to a certain value as an infinite number of the simulations are carried out. Therefore,

$$\eta = \lim_{N\to\infty} \frac{\kappa}{N} \tag{4.13}$$

where η is the ratio of the number of dissociating complexes κ to the total number of initial complexes N. Now from the principle of ESI mass spectrometric measurements of K_d, one can derive the error in K_d (ΔK_d) due to complex dissociation in the droplet regime of an ESI process. For the derivation, we have to consider the following four quantities: (1) η as defined in Equation 4.13, (2) the bulk equilibrium constant of dissociation of a protein dimer (denoted as AB):

$$K_{bulk} = \frac{[A][B]}{[AB]} = \frac{N_A N_B}{N_{AB} V} \tag{4.14}$$

where $[X]$ and N_A are the concentration and the number, respectively, of chemical species X, and V is the volume of the bulk solution used in ESI-MS, and (3) the equilibrium constant of dissociation measured in the gas-phase:

$$K_{vap}^{ideal} = \frac{N_A N_B}{N_{AB}} = K_{bulk} V \tag{4.15}$$

where the superscript *ideal* indicates that the gas-phase ensemble of protein species is completely reflective of their solution-state equilibria, and (4) the gas-phase equilibrium constant when κ dimers undergo complex dissociation:

$$K_{vap}^{real} = \frac{(N_A + \kappa)(N_B + \kappa)}{N_{AB} - \kappa} \tag{4.16}$$

where the superscript *real* indicates the actual value of K_d detected in the gas-phase when some dimers experience complex dissociation during desolvation. Using the facts that $[A] = [B]$ (thus $N_A = N_B$) and $N_{AB} = N$, one can reform Equation 4.16 to the following expression:

$$K_{vap}^{real} = \frac{K_{bulk} V + \eta V \left(2\sqrt{K_{bulk}[AB]} + \eta[AB]\right)}{1 - \eta} \tag{4.17}$$

Therefore,

$$\Delta K_d = K_{vap}^{ideal} - K_{vap}^{real} = K_{bulk} V \left[1 - \frac{1 + \eta K_{bulk}^{-1} \left(2\sqrt{K_{bulk}[AB]} + \eta[AB]\right)}{1 - \eta} \right] \tag{4.18}$$

This is the contribution from in-source complex dissociation in the droplet stage to the error in the measurement of K_d of a protein dimer by using ESI-MS. The beauty of Equation 4.18 can be found in the fact that all the variables in the expression are measurable in experiments and simulations.

4.4 Conclusion

This chapter focuses on protein complex dissociation as a result of macroion-droplet interactions. We have reported the first simulation study on how the intervening droplet environment affects the stability of a weak protein complex in ESI-generated aqueous droplets. This study is of particular significance in the application of ESI-MS for determining the equilibrium constant and the architecture of protein complexes and assemblies.

In this study, a feasible molecular dynamics set-up is suggested, which enables one to model the various pH conditions that appear in the constantly changing chemical environment of a charged droplet due to solvent evaporation (or condensation). A new methodology is also presented to find whether the complex will dissociate during the droplet evolution. The method can be used to study possible conformational changes of macroions under certain pH conditions during the lifespan of the droplet.

By performing the droplet simulations under the conditions of high temperature and low pH, it is found that the complex dissociation rates for the 2MRO and 2PEA complexes may be comparable or slower to the droplet evaporation rate. If the complex dissociation rate is considerably lower than the rate of solvent evaporation, then the dissociation would be unlikely to occur, and therefore, the association constant measured by ESI-MS would be in closer agreement with that measured in the bulk solution, provided that the complex does not undergo multiple association-dissociation events in larger droplets.

The MD simulations reveal that when complex dissociation occurs, it takes place via a multistep process. The protomers reorient and create new interfaces. At least two distinct protein-protein interfaces in the complex are identified. One of the interfaces is associated with complex destabilization. The complex first dissociates within the droplet and then its protomers move apart, elongating the droplet. Then the complete splitting of the droplet carries the two protomers away from each other. The other interface, on the other hand, results in complex stabilization after protomer reorientation.

This study also focuses on the effects of the charge distribution on the protein surface and the droplet morphology. Different protonation states may strengthen or weaken the stability of the protein-protein interface. The enhancement in the stability can be achieved by the formation of additional salt bridges which, in turn, lock the conformation of the protein complex. The destabilization of the interface, on the other hand, can arise from the electrostatic repulsion between positively charged residues located at the new interface after protomer reorientation.

The high charge state of the complex can give rise to “spiky” droplet morphologies. The rate of solvent evaporation of these shapes is higher than those of the spherical shapes. Although the rate of solvent evaporation is not directly related to the rate of protein dissociation, a higher evaporation rate assists the complex to be transferred in the gaseous state without dissociation. Also we proposed that the formation of “star”-shaped droplets is part of the mechanism that accounts for the charge state of a protein. The “spiky” droplet morphology may leave a number of amino acids bare of solvent, which creates a variable chemical environment where the amino acids are found in a droplet.

From the study of the dissociation of 2MRO and 2PEA, there are general features in the dissociation mechanism that are expected to appear in other complexes as well. These general features include (a) the role of the conical protrusions in the charge-induced instability regime in stabilizing the complex by enhancing the evaporation rate, (b) the multistep process of the complex dissociation by the formation of new interfaces with various degrees of stability, and (c) the manner that the division of the droplet occurs by carrying the protomers apart. The study provides an insight into identifying order parameters that may be used to describe the dissociation mechanism by computing the free energy along these parameters.

Lastly, we proposed a methodology that uses equilibrium MD simulations to estimate the error in the measurement of the dissociation equilibrium constant due to unwanted complex dissociation in the droplet regime. This methodology enables us to derive an expression that calculates the error in K_d, in which all the involved quantities can be obtained directly from experiments and simulations.

References

[1] Lössl, P., van de Waterbeemd, M. & Heck, A. J. The diverse and expanding role of mass spectrometry in structural and molecular biology. *EMBO J.* **35**, 2634–2657 (2016).

[2] Robinson, C. V., Sali, A. & Baumeister, W. The molecular sociology of the cell. *Nature* **450**, 973–982 (2007).

[3] Kitova, E. N., El-Hawiet, A., Schnier, P. D. & Klassen, J. S. Reliable determinations of protein-ligand interactions by direct ESI-MS measurements. Are we there yet? *J. Am. Soc. Mass Spectrom.* **23**, 431–441 (2012).

[4] Barylyuk, K. *et al.* What happens to hydrophobic interactions during transfer from the solution to the gas phase? The case of electrospray-based soft ionization methods. *J. Am. Soc. Mass. Spectrom.* **22**, 1167–1177 (2011).

[5] Erba, E. B., Barylyuk, K., Yang, Y. & Zenobi, R. Quantifying protein-protein interactions within noncovalent complexes using electrospray ionization mass spectrometry. *Anal. Chem.* **83**, 9251–9259 (2011).

[6] Gülbakan, B., Barylyuk, K. & Zenobi, R. Determination of thermodynamic and kinetic properties of biomolecules by mass spectrometry. *Curr. Opin. Biotech.* **31**, 65–72 (2015).

[7] Chingin, K., Barylyuk, K. & Chen, H. On the preservation of non-covalent protein complexes during electrospray ionization. *Phil. Trans. R. Soc. A* **374**, 20150377 (2016).

[8] Benjamin, D. R., Robinson, C. V., Hendrick, J. P., Hartl, F. U. & Dobson, C. M. Mass spectrometry of ribosomes and ribosomal subunits. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7391–7395 (1998).

[9] Tito, M. A., Tars, K., Valegard, K., Hajdu, J. & Robinson, C. V. Electrospray time-of-flight mass spectrometry of the intact MS2 virus capsid. *J. Am. Chem. Soc.* **122**, 3550–3551 (2000).

[10] Laganowsky, A. *et al.* Membrane proteins bind lipids selectively to modulate their structure and function. *Nature* **510**, 172–175 (2014).

[11] Barrera, N. P. *et al.* Mass spectrometry of membrane transporters reveals subunit stoichiometry and interactions. *Nat. Methods* **6**, 585–587 (2009).

[12] Leney, A. C. & Heck, A. J. R. Native mass spectrometry: What is in the name? *J. Am. Soc. Mass Spectrom.* **28**, 5–13 (2017).

[13] Barylyuk, K., Gülbakan, B., Xie, X. & Zenobi, R. DNA oligonucleotides: A model system with tunable binding strength to study monomer–dimer equilibria with electrospray ionization-mass spectrometry. *Anal. Chem.* **85**, 11902–11912 (2013).

[14] Hilton, G. R. & Benesch, J. L. P. Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionization mass spectrometry. *J. R. Soc. Interface* **9**, 801–816 (2012).

[15] Mehmood, S., Allison, T. M. & Robinson, C. V. Mass spectrometry of protein complexes: From origins to applications. *Annu. Rev. Phys. Chem.* **66**, 453–474 (2015).

[16] Loo, J. A. Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom. Rev.* **16**, 1–23 (1997).

[17] Erba, E. B. & Zenobi, R. Mass spectrometric studies of dissociation constants of noncovalent complexes. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **107**, 199–228 (2011).

[18] Chen, F. *et al.* Applying mass spectrometry to study non-covalent biomolecule complexes. *Mass Spectrom. Rev.* **35**, 48–70 (2016).

[19] Ngounou Wetie, A. G. *et al.* Investigation of stable and transient protein-protein interactions: Past, present, and future. *Proteomics* **13**, 538–557 (2013).

[20] Yamaguchi, K. Cold-spray ionization mass spectrometry: principle and applications. *J. Mass Spectrom.* **38**, 473–490 (2003).

[21] Yao, Y., Richards, M. R., Kitova, E. N. & Klassen, J. S. Influence of sulfolane on ESI-MS measurements of protein-ligand affinities. *J. Am. Soc. Mass. Spectrom.* **27**, 498–506 (2016).

[22] Pierce, M. M., Raman, C. & Nall, B. T. Isothermal titration calorimetry of protein-protein interactions. *Methods* **19**, 213–221 (1999).

[23] Jecklin, M. C., Schauer, S., Dumelin, C. E. & Zenobi, R. Label-free determination of protein-ligand binding constants using mass spectrometry and validation using surface plasmon resonance and isothermal titration calorimetry. *J. Mol. Recognit.* **22**, 319–329 (2009).

[24] Velazquez-Campoy, A., Leavitt, S. A. & Freire, E. *Characterization of Protein-Protein Interactions by Isothermal Titration Calorimetry*, 35–54 (Humana Press, Totowa, NJ, 2004).

[25] Robinson, C. V. *et al.* Probing the nature of noncovalent interactions by mass spectrometry. a study of protein-CoA ligand binding and assembly. *J. Am. Chem. Soc.* **118**, 8646–8653 (1996).

[26] Smith, R. D., Light-Wahl, K. J., Winger, B. E. & Loo, J. A. Preservation of non-covalent associations in electrospray ionization mass spectrometry: Multiply charged polypeptide and protein dimers. *J. Mass Spectrom.* **27**, 811–821 (1992).

[27] Heck, A. J. Native mass spectrometry: a bridge between interactomics and structural biology. *Nat. Methods* **5**, 927 (2008).

[28] Rose, R. J., Damoc, E., Denisov, E., Makarov, A. & Heck, A. J. High-sensitivity orbitrap mass analysis of intact macromolecular assemblies. *Nat. Methods* **9**, 1084–1086 (2012).

[29] Berman, H. M. *et al.* The protein data bank. *Nucleic Acids Res.* **28**, 235–242 (2000).

[30] `http://www.rcsb.org`.

[31] Nowicka, U. *et al.* DNA-damage-inducible 1 protein (Ddi1) contains an uncharacteristic ubiquitin-like domain that binds ubiquitin. *Structure* **23**, 542–557 (2015).

[32] Ryabov, Y. & Fushman, D. Structural assembly of multidomain proteins and protein complexes guided by the overall rotational diffusion tensor. *J. Am. Chem. Soc.* **129**, 7894–7902 (2007).

[33] Liu, Z. *et al.* Noncovalent dimerization of ubiquitin. *Angew. Chem. Int. Ed.* **51**, 469–472 (2012).

[34] Wilber, A. W., Doye, J. P. K., Louis, A. A. & Lewis, A. C. F. Monodisperse self-assembly in a model with protein-like interactions. *J. Chem. Phys.* **131**, 175102 (2009).

[35] Klein, H. C. R. & Schwarz, U. S. Studying protein assembly with reversible brownian dynamics of patchy particles. *J. Chem. Phys.* **140**, 184112 (2014).

[36] Villar, G. *et al.* Self-assembly and evolution of homomeric protein complexes. *Phys. Rev. Lett.* **102**, 118106 (2009).

[37] McManus, J. J., Charbonneau, P., Zaccarelli, E. & Asherie, N. The physics of protein self-assembly. *Curr. Opin. Colloid Interface Sci.* **22**, 73–79 (2016).

[38] Jackson, S. E. Ubiquitin: a small protein folding paradigm. *Org. Biomol. Chem.* **4**, 1845–1853 (2006).

[39] Petsko, G. A. & Ringe, D. *Protein Structure and Function.* Primers in biology (New Science Press Sunderland, MA, London, 2004).

[40] Nooren, I. M. & Thornton, J. M. Diversity of protein-protein interactions. *EMBO J.* **22**, 3486–3492 (2003).

[41] Acuner Ozbabacan, S. E., Engin, H. B., Gursoy, A. & Keskin, O. Transient Protein-Protein Interactions. *Protein Eng. Des. Sel.* **24**, 635–648 (2011).

[42] Malevanets, A. & Consta, S. Variation of droplet acidity during evaporation. *J. Chem. Phys.* **138**, 184312 (2013).

[43] Rayleigh, L. XX. on the equilibrium of liquid conducting masses charged with electricity. *Philos. Mag.* **14**, 184–186 (1882).

[44] Hendricks, C. & Schneider, J. Stability of a conducting droplet under the influence of surface tension and electrostatic forces. *Am. J. Phys.* **31**, 450–453 (1963).

[45] Consta, S. & Malevanets, A. Disintegration mechanisms of charged nanodroplets: novel systems for applying methods of activated processes. *Mol. Simulat.* **41**, 73–85 (2015).

[46] Consta, S., Oh, M. I. & Malevanets, A. New mechanisms of macroion-induced disintegration of charged droplets. *Chem. Phys. Lett.* **663**, 1–12 (2016).

[47] Burgot, J. *Ionic Equilibria in Analytical Chemistry* (Springer New York, 2012).

[48] Ho, J. & Coote, M. A universal approach for continuum solvent pKa calculations: Are we there yet? *Theor. Chem. Acc.* **125**, 3–21 (2009).

[49] Anandakrishnan, R., Aguilar, B. & Onufriev, A. V. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* **40**, W537–W541 (2012).

[50] Antosiewicz, J., McCammon, J. & Gilson, M. The determinants of pKas in proteins. *Biochemistry (Moscow)* **35**, 7819–7833 (1996).

[51] Baptista, A., Teixeira, V. & Soares, C. Constant-pH molecular dynamics using stochastic titration. *J. Chem. Phys.* **117**, 4184–4200 (2002).

[52] Bas, D., Rogers, D. & Jensen, J. Very fast prediction and rationalization of pKa values for protein-ligand complexes. *Proteins: Structure, Function and Genetics* **73**, 765–783 (2008).

[53] Olsson, M., SØndergaard, C., Rostkowski, M. & Jensen, J. PROPKA3: Consistent treatment of internal and surface residues in empirical pKa predictions. *J. Chem. Theory Comput.* **7**, 525–537 (2011).

[54] Chen, J., Brooks III, C. & Khandogin, J. Recent advances in implicit solvent-based methods for biomolecular simulations. *Curr. Opin. Struct. Biol.* **18**, 140–148 (2008).

[55] Cramer, C. & Truhlar, D. A universal approach to solvation modeling. *Acc. Chem. Res.* **41**, 760–768 (2008).

[56] Gilson, M. Theory of electrostatic interactions in macromolecules. *Curr. Opin. Struct. Biol.* **5**, 216–223 (1995).

[57] Khandogin, J. & Brooks III, C. Constant pH molecular dynamics with proton tautomerism. *Biophys. J.* **89**, 141–157 (2005).

[58] Schreiber, G., Haran, G. & Zhou, H. X. Fundamental aspects of protein-protein association kinetics. *Chem. Rev.* **109**, 839–860 (2009).

[59] Warshel, A., Sharma, P., Kato, M. & Parson, W. Modeling electrostatic effects in proteins. *Biochim. Biophys. Acta, Proteins Proteomics* **1764**, 1647–1676 (2006).

[60] Grimsley, G. R., Scholtz, J. M. & Pace, C. N. A summary of the measured pK values of the ionizable groups in folded proteins. *Protein Sci.* **18**, 247–251 (2009).

[61] Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comp. Chem.* **26**, 1781–1802 (2005).

[62] Koopman, E. A. & Lowe, C. P. Advantages of a Lowe-Andersen thermostat in molecular dynamics simulations. *J. Chem. Phys.* **124**, 204103 (2006).

[63] Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comp. Chem.* **31**, 671–690 (2010).

[64] Yu, W., He, X., Vanommeslaeghe, K. & MacKerell, A. D. Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations. *J. Comp. Chem.* **33**, 2451–2468 (2012).

[65] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

[66] Jorgensen, W. L. & Jenson, C. Temperature dependence of TIP3P, SPC, and TIP4P water from NPT monte carlo simulations: Seeking temperatures of maximum density. *J. Comp. Chem.* **19**, 1179–1186 (1998).

[67] Vega, C. & de Miguel, E. Surface tension of the most popular models of water by using the test-area simulation method. *J. Chem. Phys.* **126**, 154707 (2007).

[68] Chen, F. & Smith, P. E. Simulated surface tensions of common water models. *J. Chem. Phys.* **126**, 221101 (2007).

[69] Vega, C. & Abascal, J. L. F. Simulating water with rigid non-polarizable models: a general perspective. *Phys. Chem. Chem. Phys.* **13**, 19663–19688 (2011).

[70] Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **14**, 33–38 (1996).

[71] Soleilhac, A., Dagany, X., Dugourd, P., Girod, M. & Antoine, R. Correlating droplet size with temperature changes in electrospray source by optical methods. *Anal. Chem.* **87**, 8210–8217 (2015).

[72] Gibson, S. C., Feigerle, C. S. & Cook, K. D. Fluorometric measurement and modeling of droplet temperature changes in an electrospray plume. *Anal. Chem.* **86**, 464–472 (2013).

[73] Servage, K. A., Silveira, J. A., Fort, K. L., Clemmer, D. E. & Russell, D. H. Water-mediated dimerization of ubiquitin ions captured by cryogenic ion mobility-mass spectrometry. *J. Phys. Chem. Lett.* **6**, 4947–4951 (2015).

[74] Young, L., Jernigan, R. & Covell, D. A role for surface hydrophobicity in protein-protein recognition. *Protein Sci.* **3**, 717–729 (1994).

[75] Dikic, I., Wakatsuki, S. & Walters, K. J. Ubiquitin-binding domains - from structures to functions. *Nat. Rev. Mol. Cell. Biol.* **10**, 659–671 (2009).

[76] Mueller, T. D. & Feigon, J. Solution structures of UBA domains reveal a conserved hydrophobic surface for protein-protein interactions. *J. Mol. Biol.* **319**, 1243–1255 (2002).

[77] Hurley, J. H., Lee, S. & Prag, G. Ubiquitin-binding domains. *Biochem. J* **399**, 361–372 (2006).

[78] Darnell, S. J., Page, D. & Mitchell, J. C. An automated decision-tree approach to predicting protein interaction hot spots. *Proteins: Structure, Function, and Bioinformatics* **68**, 813–823 (2007).

[79] Zhu, X. & Mitchell, J. C. KFC2: A knowledge-based hot spot prediction method based on interface solvation, atomic density, and plasticity features. *Proteins: Struct., Funct., Bioinf.* **79**, 2671–2683 (2011).

[80] Moreira, I. S., Fernandes, P. A. & Ramos, M. J. Hot spots - a review of the protein-protein interface determinant amino acid residues. *Proteins: Struct., Funct., Bioinf.* **68**, 803–812 (2007).

[81] Lehmann, E. & Casella, G. *Theory of Point Estimation* (Springer Verlag, 1998).

[82] Schilder, J. & Ubbink, M. Formation of transient protein complexes. *Curr. Opin. Struct. Biol.* **23**, 911–918 (2013).

[83] Sinz, A., Arlt, C., Chorev, D. & Sharon, M. Chemical cross-linking and native mass spectrometry: A fruitful combination for structural biology. *Protein Sci.* **24**, 1193–1209 (2015).

[84] Kastritis, P. L. & Bonvin, A. M. J. J. On the binding affinity of macromolecular interactions: Daring to ask why proteins interact. *J. R. Soc. Interface* **10**, 20120835 (2012).

[85] Consta, S. Fragmentation reactions of charged aqueous clusters. *Comp. Theor. Chem.* **591**, 131–140 (2002).

[86] Consta, S., Mainer, K. R. & Novak, W. Fragmentation mechanisms of aqueous clusters charged with ions. *J. Chem. Phys.* **119**, 10125–10132 (2003).

[87] Dole, M. *et al.* Molecular beams of macroions. *J. Chem. Phys.* **49**, 2240–2249 (1968).

[88] Nguyen, S. & Fenn, J. B. Gas-phase ions of solute species from charged droplets of solutions. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 1111–1117 (2007).

[89] de la Mora, J. F. Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Anal. Chim. Acta* **406**, 93–104 (2000).

[90] Felitsyn, N., Peschke, M. & Kebarle, P. Origin and number of charges observed on multiply-protonated native proteins produced by ESI. *Int. J. of Mass Spectrom.* **219**, 39–62 (2002).

[91] Peschke, M., Blades, A. & Kebarle, P. Charged states of proteins. reactions of doubly protonated alkyldiamines with NH_3: Solvation or deprotonation. Extension of two proton cases to multiply protonated globular proteins observed in the gas phase. *J. Am. Chem. Soc.* **124**, 11519–11530 (2002).

[92] Heck, A. J. & van den Heuvel, R. H. Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom. Rev.* **23**, 368–389 (2004).

[93] Catalina, M. I., van den Heuvel, R. H., van Duijn, E. & Heck, A. J. Decharging of globular proteins and protein complexes in electrospray. *Chem. Eur. J.* **11**, 960–968 (2005).

[94] Consta, S. Manifestation of Rayleigh instability in droplets containing multiply charged macroions. *J. Phys. Chem. B* **114**, 5263–5268 (2010).

[95] Jungwirth, P. & Cremer, P. S. Beyond Hofmeister. *Nat. Chem.* **6**, 261–263 (2014).

[96] Han, L., Hyung, S.-J. & Ruotolo, B. T. Dramatically stabilizing multiprotein complex structure in the absence of bulk water using tuned Hofmeister salts. *Faraday Discuss.* **160**, 371–388 (2013).

[97] Hofmeister, F. Zur lehre von der wirkung der salze. *Arch. Exp. Pathol. Parmakol. (Leipzig)* **24**, 247–260 (1888).

[98] Konermann, L. Addressing a common misconception: Ammonium acetate as neutral pH "buffer" for native electrospray mass spectrometry. *J. Am. Soc. Mass Spectrom.* **28**, 1827–1835 (2017).

Chapter 5

Conclusion

This dissertation probed in a systematic manner the onset and various outcomes of macroion-droplet interactions that involve a single macromolecule or macromolecular complex in a nanoscopic droplet. As a direct consequence of macroion-droplet interactions, the droplet systems undergo distinct disintegration mechanisms and shape deformations. The three general classes of the droplet-macroion interactions that have been identified in the Consta group are: (a) the extrusion of a linear macroion from a droplet, (b) the “pearl-necklace” droplet conformation, and (c) the “star”-shaped droplets. These classes are universal for macroions of different levels of complexity ranging from simple macroions to assemblies of proteins.

By using atomistic molecular dynamics, I identified new motifs of charge-induced droplet disintegration and deformations that fall into the general classes. Analytical models were then formulated for those mechanisms. Based on the analytical theory and the simulation data, suggestions were made to resolve critical issues in electrospray ionization mass spectrometry (ESI-MS) and its relevant research fields.

Extrusion of a linear macroion from a droplet has been evidenced for poly(ethylene glycol) by atomistic simulations and experiments [1–4]. An analytical theory that considers the solvation energy of a macroion in a droplet and its electrostatic energy enables one to identify extrusion mechanisms [5, 6] and test them using simulations of a model linear macroion. I studied systematically the charging and release mechanisms of a flexible macromolecule, modelled by poly(ethylene glycol) (PEG), in a charged droplet of water, acetonitrile, and their mixtures. I compared how PEG is solvated and charged by sodium (Na^+) ions in a droplet. It was found that the location of the macromolecule and the droplet size affect the dimension of PEG conformation. PEG is charged on the surface of a sodiated aqueous droplet that is found close to the Rayleigh limit. Its charging is coupled to the extrusion mechanism, where PEG segments leave the droplet once they coordinate a Na^+ ion or in a correlated motion with Na^+ ions. Here, this macroion extrusion was named the Charging-Induced Extrusion Mechanism (CI-EM). We then used the simulation data to construct an analytical model for the CI-EM, which suggests that the droplet surface electric field may play a role in the macroion-droplet interactions that lead to the extrusion of the macroion. This study provides the first evidence of the effect of the surface electric field by using atomistic simulations. In contrast, as PEG resides in the interior of an acetonitrile droplet, it is sodiated inside the droplet. The compact macroion transitions through partially unwound states to an extended conformation, a process occurring during the final stage of desolvation and in the presence of only a handful of acetonirile molecules. For charged water/acetonitrile droplets, the sodiation of PEG is determined by the aqueous component, reflecting its slower evaporation and preference over acetonitrile for solvating Na^+ ions. This study shows that the final charge state of a flexible macromolecule is determined not merely by the solvent surface tension, but by unique macromolecule-ion-solvent interactions. This has not been clarified in ESI experiments due to the lack of molecular-level understanding of the mechanism, giving rising to controversy over the role of solvent in determining the macroion charge state [7–13].

In this study, however, only Na^+ ions were used as charge carriers that interact with the flexible macromolecule. In the modelling of the charging of PEG, Consta and her co-workers have performed molecular simulations with lithium and calcium ions [4, 14]. These cations give the same extrusion mechanism of the sodiated PEG from an aqueous droplet. This consistency indicates that variations of interactions within a certain range do not alter the observed mechanism [15]. Therefore, it is an engaging question whether this consistency is also observed for polyatomic ions. The nature of charge carriers may have a striking impact on how a flexible macromolecule is released from the droplet. In particular, the size and the molecular geometry of charge carriers may affect charging of the macromolecule. For example, large, complex ions may require a longer sequence of monomers when they are looped by the macromolecule, possibly reducing the final charge state of the macroion.

The next question I addressed is whether a protein that lacks an ordered three-dimensional structure extrudes from an aqueous droplet. Extrusion of a linear macroion from an aqueous droplet has been evidenced for a charged poly(ethylene glycol) but the question for proteins following the same mechanism of extrusion still remains unresolved. Therefore, I simulated examples of proteins drawn from the broad classes of membrane proteins with ordered and disordered domains, intrinsically disordered proteins, and denatured proteins. It was found that an unstructured protein does not extrude from an aqueous nanoscopic droplet. Instead, the water accumulates in the two termini of the protein chain forming two "pear-shaped" lobes. The intermediate segment of the extended protein chain is sparsely solvated. This "pearl-necklace" droplet conformation may emerge due to the interplay of a number of factors, which are (i) the tendency of the solvent to form spherical droplets in order to minimize the surface energy, (ii) the constraint that the charge of each sub-droplet ("pearl") should be below the Rayleigh limit, (iii) the solvation energy of the chain, (iv) the length of the chain and the distribution of charge. The water may continue to evaporate from the chain termini giving rise to charge-induced instability where the solvent forms conical extrusions, which resemble multi-point "stars" . As a sub-droplet may also transform into a "star", there may be more than one "star"-shaped regions along the chain. I also discussed that the "spiky" structures may affect the charge state of a macroion.

A charged spherical droplet composed of a solvent and a non-fissile macroion (spherical or linear) above its Rayleigh limit deforms into a stable "star"-like shape. Proteins and nucleic acids are examples of macroions that may give rise to the spiky shapes [16]. I focused on a dielectric droplet that contains a single spherical macroion located at the centre of the droplet. As all the charges are bound to the central macroion, droplet fissions are not possible in the charge-induced instability regime. Instead, spherical dielectric droplets are deformed into "star" shapes when the droplet charge-squared-to-volume ratio exceeds a well-defined threshold value. Here I analyzed the geometry of the star morphologies and its relation to the net charge of the droplet, and compared droplets composed of dimethyl sulfoxide (DMSO) and water. It was revealed that the theoretical and simulation results for the onset of instability are in excellent agreement. The same charge-squared-to-volume ratio in the two solvents lead to slightly different numbers of rays, but over a wide range of the droplet charges, there exist number of points in the stars that are less preferred. Four point-stars in particular are highly preferred in both water and DMSO. Positive and negative ions show the same number of rays in both water and DMSO. I presented an analytical description of the number of rays as a function of the charge that shows a quadratic relation, which is in close agreement with the trends found in the simulations. By analyzing the angle of the rays, it was found that they do not follow the Taylor cone predictions of 100° for aqueous droplets, even though the amount of solvent is sufficient to arrange in a structure that satisfies the 100°.

Regarding the intriguing finding of "star"-shaped droplets, there are still a number of questions to be answered. One needs to establish the relation between the number of points of the star-shaped droplets to the Thomson problem [17] of the distribution of electrons in a spherical shell. The relation between the angle of the Taylor cone [18] to the observed angles of the spikes is yet to be investigated.

As the next level of system complexity, I studied a weak transient protein complex in a charged aqueous droplet. Maintaining the interface of a weak transient protein complex transferred from bulk solution to the gaseous state via evaporating droplets is a critical question in the detection of the complex dissociation constant by using ESI-MS. I explored the factors that may affect the stability of the protein-protein interaction (PPI), using the complex of ubiquitin and the ubiquitin-associated domain and the noncovalent complex of diubiquitin in aqueous droplets. A general method was presented to determine the protonation states of the complexes we investigated in particular, and that of a protein in general, at various pH conditions that an evaporating droplet acquires due to its change in size. It was found that the combination of high temperature and high charge states of the protein complexes may destabilize the interface by creating new interfaces instead of a direct rupture of the initial stable interface. My simulations provided evidence that highly charged protein complexes are found in droplets that form conical extrusions of the solvent on the surface due to charge-induced instability. This distinct droplet morphology leads to a higher solvent evaporation rate that assists in transferring the complex in the gaseous state without dissociation. Also, I discussed that the formation of spikes may provide certain pathways for ion release. The "star" morphologies may delay the release of protons from the droplet, and thus, they may affect the charge state of an electrosprayed protein. The conical solvent protrusions can also expose to the vapour phase certain amino acids that otherwise would be solvated in a droplet with the protein complex of low charge states. The new vapour-protein interface does not have a direct effect on the stability of the PPI. Finally, I proposed a method that uses equilibrium direct molecular dynamics to compute complex dissociation rates and so the error in the dissociation constant measurement due to complex dissociation in the droplet.

This research will be complemented by additional studies on the effect of different co-ions and counter-ions on the stabilization of PPIs in the droplet regime. The effect of hydration of salt ions on salting-out of biological macromolecules and colloidal particles was first studied by Franz Hofmeister in 1888 [19]. The so-called Hofmeister series is a classification of ions in order of their ability to salt out proteins in solution. Later, ion binding to biological macromolecules and its influence on their conformational preference has been widely studied in experiments and simulations [20–25]. The interactions of ions in biological systems have been related to their hydration properties [19, 26–28]. Ruotolo and his co-workers [24, 29] used ion mobility-mass spectrometry to examine the influence of bound cations and anions on the stability of multi-protein complexes in the absence of bulk solvent. They ascribed the difference between their rank order of the stabilizing ions and the Hofmeister series to the lack of protein solvation and ion hydration in the gas phase [24, 29]. However, little is known about the role of various cations and anions in the stabilization of macromolecular complex ions in the droplet regime. Hence, this is the next question to be explored.

The studies on charged droplets involving macromolecules are important in several research fields. Since a droplet is a confining environment for chemical species and their reactions, the studies play a pivotal role in understanding physical chemistry in small volumes (see Appendix D as an example). A significant application is chemical reactions in nanofluidic environments [30, 31]. Some of the distinct droplet morphologies that have been identified may have potential applications in materials science and catalysis. Another major application of the studies is found in the mass spectrometry. Liquid charged droplets that carry macroions from bulk solution to the gaseous phase for mass spectrometry analysis are intermediate states in many ionization techniques [32–36], among which are electrospray [3, 37], sonic spray [38], and thermospray [39] to mention but a few. The origin and the magnitude of the final macroion charge state is determined to a large extent by the droplet chemistry. Therefore, understanding the droplet chemistry at the molecular level is necessary in the field of mass spectrometry.

References

[1] de la Mora, J. F. Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Anal. Chim. Acta* **406**, 93–104 (2000).

[2] Nohmi, T. & Fenn, J. B. Electrospray mass spectrometry of poly(ethylene glycols) with molecular weights up to five million. *J. Am. Chem. Soc.* **114**, 3241–3246 (1992).

[3] Wong, S. F., Meng, C. K. & Fenn, J. B. Multlple charging in electrospray ionization of poly(ethylene glycols). *J. Chem. Phys.* **92**, 546–550 (1988).

[4] Consta, S. & Chung, J. K. Charge-induced conformational changes of PEG-(Na_n^+) in a vacuum and aqueous nanodroplets. *J. Phys. Chem. B* **115**, 10447–10455 (2011).

[5] Consta, S. & Malevanets, A. Manifestations of charge induced instability in droplets effected by charged macromolecules. *Phys. Rev. Lett.* **109**, 148301 (2012).

[6] Consta, S., Oh, M. I. & Malevanets, A. New mechanisms of macroion-induced disintegration of charged droplets. *Chem. Phys. Lett.* **663**, 1–12 (2016).

[7] Sterling, H. J. & Williams, E. R. Origin of supercharging in electrospray ionization of noncovalent complexes from aqueous solution. *J. Am. Soc. Mass Spectrom.* **20**, 1933–1943 (2009).

[8] Šamalikova, M. & Grandori, R. Protein charge-state distributions in electrospray-ionization mass spectrometry do not appear to be limited by the surface tension of the solvent. *J. Am. Chem. Soc.* **125**, 13352–13353 (2003).

[9] Brahim, B., Alves, S., Cole, R. B. & Tabet, J.-C. Charge enhancement of single-stranded dna in negative electrospray ionization using the supercharging reagent meta-nitrobenzyl alcohol. *J. Am. Soc. Mass Spectrom.* **24**, 1988–1996 (2013).

[10] Šamalikova, M. & Grandori, R. Testing the role of solvent surface tension in protein ionization by electrospray. *J. Mass Spectrom.* **40**, 503–510 (2005).

[11] Lomeli, S. H., Yin, S., Ogorzalek Loo, R. R. & Loo, J. A. Increasing charge while preserving noncovalent protein complexes for ESI-MS. *J. Am. Soc. Mass Spectrom.* **20**, 593–596 (2009).

[12] Lomeli, S. H., Peng, I. X., Yin, S., Ogorzalek Loo, R. R. & Loo, J. A. New reagents for increasing ESI multiple charging of proteins and protein complexes. *J. Am. Soc. Mass Spectrom.* **21**, 127–131 (2010).

[13] Ogorzalek Loo, R. R., Lakshmanan, R. & Loo, J. A. What protein charging (and supercharging) reveal about the mechanism of electrospray ionization. *J. Am. Soc. Mass Spectrom.* **25**, 1675–1693 (2014).

[14] Chung, J. K. & Consta, S. Release mechanisms of poly(ethylene glycol) macroions from aqueous charged nanodroplets. *J. Phys. Chem. B* **116**, 5777–5785 (2012).

[15] Consta, S., Oh, M. I. & Soltani, S. Advances in the theoretical and molecular simulation studies of the ion chemistry in droplets. *Int. J. Mass Spectrom.* **377**, 557–567 (2015).

[16] Sharawy, M. & Consta, S. Characterization of “star” droplet morphologies induced by charged macromolecules. *J. Phys. Chem. A* **120**, 8871–8880 (2016).

[17] Thomson, J. XXIV. on the structure of the atom: an investigation of the stability and periods of oscillation of a number of corpuscles arranged at equal intervals around the circumference of a circle; with application of the results to the theory of atomic structure. *Philos. Mag. Ser. 6* **7**, 237–265 (1904).

[18] Taylor, G. Disintegration of water drops in an electric field. *Proc. Royal Soc. A* **280**, 383–397 (1964).

[19] Hofmeister, F. Zur lehre von der wirkung der salze. *Arch. Exp. Pathol. Parmakol. (Leipzig)* **24**, 247–260 (1888).

[20] Chen, L., Gao, Y. Q. & Russell, D. H. How alkali metal ion binding alters the conformation preferences of Gramicidin A: A molecular dynamics and ion mobility study. *J. Phys. Chem. A* **116**, 689–696 (2012).

[21] Flick, T. G., Merenbloom, S. I. & Williams, E. R. Effects of metal ion adduction on the gas-phase conformations of protein ions. *J. Am. Soc. Mass Spectrom.* **24**, 1654–1662 (2013).

[22] Das, R. K. & Pappu, R. V. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 13392–13397 (2013).

[23] Makhatadze, G. I., Lopez, M. M., Richardson, J. M. & Thmos, S. T. Anion binding to the ubiquitin molecule. *Protein Sci.* **7**, 689–697 (1998).

[24] Han, L., Hyung, S.-J. & Ruotolo, B. T. Dramatically stabilizing multiprotein complex structure in the absence of bulk water using tuned Hofmeister salts. *Faraday Discuss.* **160**, 371–388 (2013).

[25] Wagner, N. D., Kim, D. & Russell, D. H. Increasing ubiquitin ion resistance to unfolding in the gas phase using chloride adduction: Preserving more ”native-like” conformations despite collisional activation. *Anal. Chem.* **88**, 5934–5940 (2016).

[26] Collins, K. Sticky ions in biological systems. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5553–5557 (1995).

[27] Washabaugh, M. & Collins, K. The systematic characterization by aqueous column chromatography of solutes which affect protein stability. *J. Biol. Chem.* **261**, 12477–12485 (1986).

[28] Jungwirth, P. & Cremer, P. S. Beyond Hofmeister. *Nat. Chem.* **6**, 261–263 (2014).

[29] Han, L., Hyung, S.-J., Mayers, J. J. S. & Ruotolo, B. T. Bound anions differentially stabilize multiprotein complexes in the absence of bulk solvent. *J. Am. Chem. Soc.* **133**, 11358–11367 (2011).

[30] Song, H., Chen, D. L. & Ismagilov, R. F. Reactions in droplets in microfluidic channels. *Angew. Chem. Int. Ed.* **45**, 7336–7356 (2006).

[31] Song, Y. *et al.* Fabrication of fibrillosomes from droplets stabilized by protein nanofibrils at all-aqueous interfaces. *Nat. Commun.* **7**, 12934 (2016).

[32] Monge, M. E., Harris, G. A., Dwivedi, P. & Fernández, F. M. Mass spectrometry: recent advances in direct open air surface sampling/ionization. *Chem. Rev.* **113**, 2269–2308 (2013).

[33] Teunissen, S. F. & Eberlin, M. N. Transferring ions from solution to the gas phase: The two basic principles. *J. Am. Soc. Mass Spectrom.* **28**, 2255–2261 (2017).

[34] Sakairi, M. & Kambara, H. Atmospheric pressure spray ionization for liquid chromatography/mass spectrometry. *Anal. Chem.* **61**, 1159–1164 (1989).

[35] Bruins, A. P., Covey, T. R. & Henion, J. D. Ion spray interface for combined liquid chromatography/atmospheric pressure ionization mass spectrometry. *Anal. Chem.* **59**, 2642–2646 (1987).

[36] Kambara, H. Sample introduction system for atmospheric pressure ionization mass spectrometry of nonvolatile compounds. *Anal. Chem.* **54**, 143–146 (1982).

[37] Dole, M. *et al.* Molecular beams of macroions. *J. Chem. Phys.* **49**, 2240–2249 (1968).

[38] Hirabayashi, A., Sakairi, M. & Koizumi, H. Sonic spray ionization method for atmospheric pressure ionization mass spectrometry. *Anal. Chem.* **66**, 4557–4559 (1994).

[39] Blakley, C. R. & Vestal, M. L. Thermospray interface for liquid chromatography/mass spectrometry. *Anal. Chem.* **55**, 750–754 (1983).

Appendix A

Poly(ethylene glycol) in Neutral Droplets

A.1 Modelling and Computational Methods

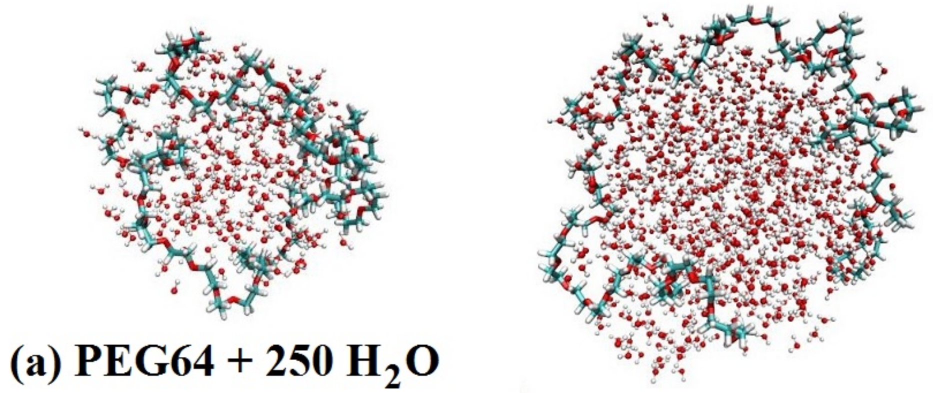
PEG64 was solvated in H_2O and MeCN droplets with varying sizes (250-7000 H_2O molecules and 250-1500 MeCN molecules). Equilibrium runs were performed by placing the droplet system into a cubic simulation box with the periodic boundary condition (PBC). The size of the simulation box was approximately three to four times larger than the diameter of the droplet so as to ensure that the system does not interact with its images and to allow for the shape fluctuations of the droplets. The temperature was set at 300 K and 250 K for water and MeCN droplets, respectively, so that vigorous solvent evaporation was suppressed. The time step of integration was set to 1 fs in all the runs. The long-distance electrostatic and van der Waals interactions between non-bonded atoms were truncated using the switch function. In this scheme, the Coulomb and LJ potential functions are decreased over the entire range, and the forces decay smoothly to zero between the switch and the cutoff radius, so that the forces and their derivatives are continuous at the cut-off radius. The switch distances employed for 250 H_2O, 800 H_2O, 1500 H_2O, 4000 H_2O, and 7000 H_2O are 4.0 nm, 5.0 nm, 5.5 nm, 7.5 nm, and 9.0 nm, respectively. Those for 250 MeCN, 800 MeCN, and 1500 MeCN are 4.5 nm, 6.5 nm, and 8.0 nm, respectively. The cut-off radii are 1 nm longer than the corresponding switch distances.

The simulations were prepared with different initial conditions where either a compact or stretched conformation of PEG64 was initially placed inside the droplets for the aqueous droplets with more than 250 H_2O molecules. Furthermore, four additional aqueous droplets consisting of 4000 water molecules, and PEG with 30, 40, 50 and 64 monomers were simulated at the same conditions as described above for 4 ns, so as to test the location of PEG in the droplets. In these systems, a compact PEG was initially solvated inside the aqueous droplets. A simulation of PEG64 *in vacuo* without any solvent was also performed for 100 ns for the purpose of comparisons with the other systems.

A.2 PEG on the Surface of an Aqueous Droplet

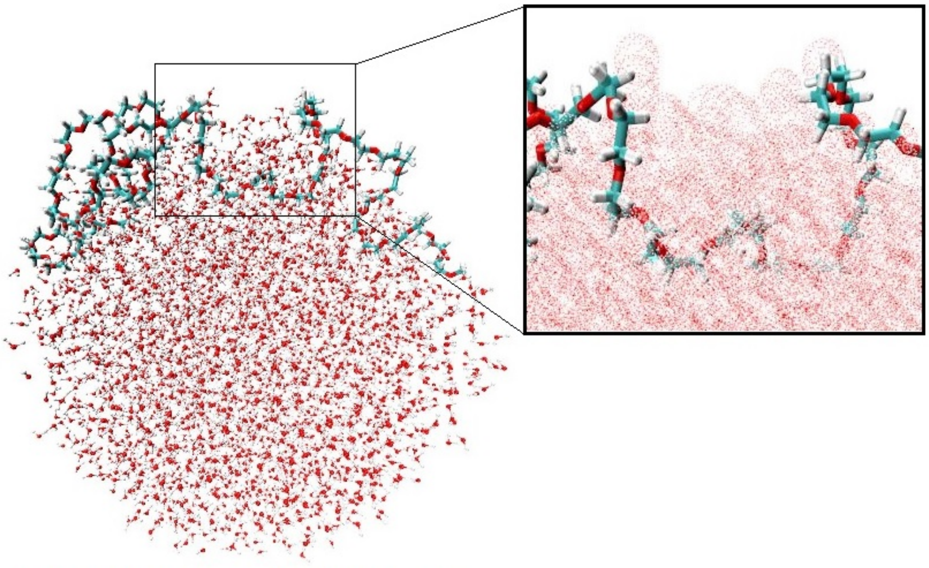
Typical snapshots of the solvation of PEG64 in aqueous droplets of various sizes are shown in Figure A.1 (a-c). In all the systems at equilibrium, PEG consistently resides on the surface of the droplets. The statistical sampling of the PEG location was tested by preparing ten initial configurations where the PEG64 was well solvated in the interior of the droplet. Among these configurations, droplets of 4000 water molecules with PEG30, PEG40, PEG50, and PEG64 were simulated, starting from PEG at the centre of the droplet. In all the systems, the PEG phase-separated from the aqueous droplet within 4 ns. We found that there was no obvious correlation between the length of PEG and the diffusion time. It is interesting that the PEG64 does not entirely adhere to the surface of the droplet; rather, it often forms multiple loops that alternate between regions immersed into the solvent and regions detached from the water/vapour interface of the droplet. These alternating loops are shown in the inset of Figure A.1 (c). This partial burial of the macromolecule is common regardless of the size of the aqueous droplet, as shown in Figure A.2 which shows the probability density (P) finding a certain distance between the ethereal oxygen atoms of PEG and the centre of mass of the droplet. The peak corresponds to the radius of the droplet; that is, the oxygen sites adhere to the droplet surface. However, the distribution implies that the oxygen atoms may penetrate deep inside the droplet and detach from the droplet surface due to the formation of the alternating loops. The loops are dynamic and the time-scale that a segment of PEG64 is immersed into the droplet may vary from several tens of picoseconds to a few hundreds of picoseconds. The alternating loops of PEG have been consistently found at the water/vapour interface in modeling performed by Prasitnok *et al.* [1]. In those studies, coarse-grained modeling of multiple PEG chains was performed in the liquid/vapour interface of an aqueous solution.

The conformations of PEG64 were quantified by measuring the radius of gyration (R_g) and



(a) PEG64 + 250 H_2O

(b) PEG64 + 800 H_2O



(c) PEG64 + 4000 H_2O

Figure A.1: Typical snapshots (a-c) of aqueous droplets of various sizes that contain a PEG64 molecule. The inset of (c) shows the magnification of a region of (c) that shows the alternating loops in and out of an aqueous droplet of 4000 H_2O molecules. The H_2O molecules are represented by dotted red spheres.

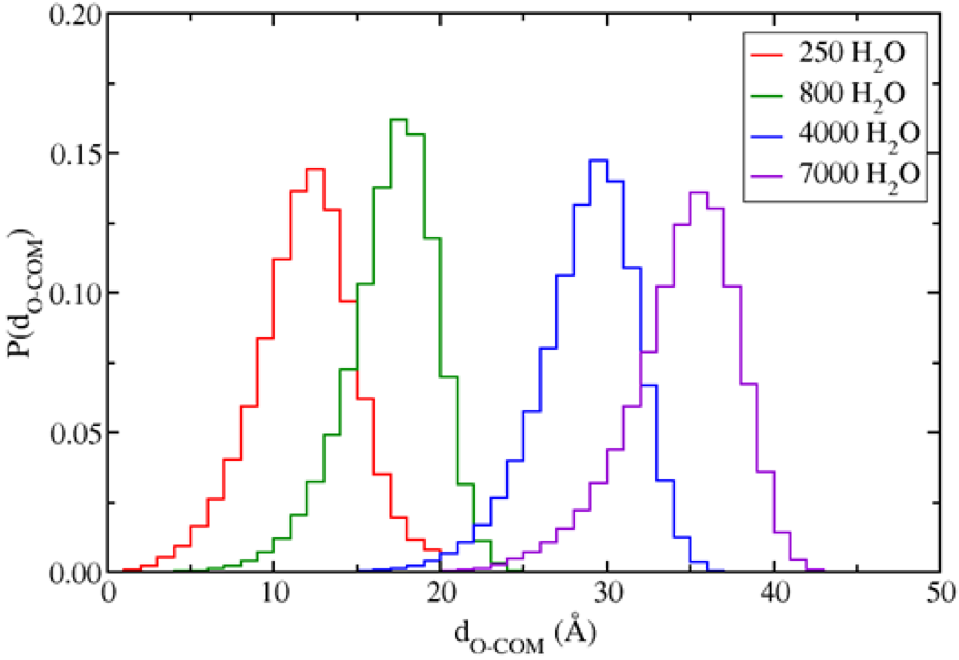


Figure A.2: Probability density (*P*) of the distance between the ethereal oxygen atoms of PEG and the centre of mass of the aqueous droplet. Different colours indicate different sizes of H_2O molecules: red for 250 H_2O, green for 800 H_2O, blue for 4000 H_2O, and purple for 7000 H_2O.

the end-to-end distance (R_{EE}) of the polymer. The histograms of the R_g and R_{EE} distributions are shown in Figures A.3 (b) and A.4 (a). The most probable values of R_g and R_{EE} converge to an upper limit as the droplet size increases. The convergence of the most probable values of R_g and R_{EE} with increasing droplet size can be also demonstrated by plotting these values *vs.* the number of H_2O molecules (N_{H_2O}) in a droplet (Figure A.4 (b)). This limit should be close to the chain dimensions observed in the case of a bulk water/vapour interface.

To find the origin of the convergence of the R_g and R_{EE} distributions, the maximum number of solvent molecules in the first first solvation shell of PEG64 and the droplet curvature (κ) *vs.* the number of the H_2O molecules in the droplet were examined. Figure A.3 (a) shows a rapid increase in the number of water molecules in the first solvation shell of PEG64. This maximum number is attained when the number of the solvent molecules is 1500 or more. This convergence indicates that the maximum solvent-monomer interaction sites have been saturated for the polymer at that location; in other words, there is an upper limit of the contact area available for the interaction with the solvent. As shown in the inset of Figure A.3 (a), assuming that the droplet is in a spherical shape, the curvature of the great circle decreases and converges to a value close to zero in the same manner as the number of water molecules in the first solvation shell of PEG converges. When the diameter of the droplet is relatively large ($\approx$ 7 nm), the curvature comes close to zero and the droplet surface mimics the flat surface of a solvent/vapour interface. The change of the slope in the curvature of the droplet and the maximum interaction contacts of PEG64 with H_2O *vs.* the droplet size are in agreement with that of the R_g and R_{EE} values. All the quantities show the onset of the convergence at approximately 1500 H_2O molecules (which corresponds to a droplet with diameter D of $\approx$ 4.7 nm), and they clearly reach a plateau at approximately 4000 H_2O molecules ($D \approx 6.9$ nm). The convergence of these quantities with the size of the droplet is greatly affected by the curvature of the droplet because of the location of the PEG on the surface and its specific interactions with the solvent at the droplet/vapour interface.

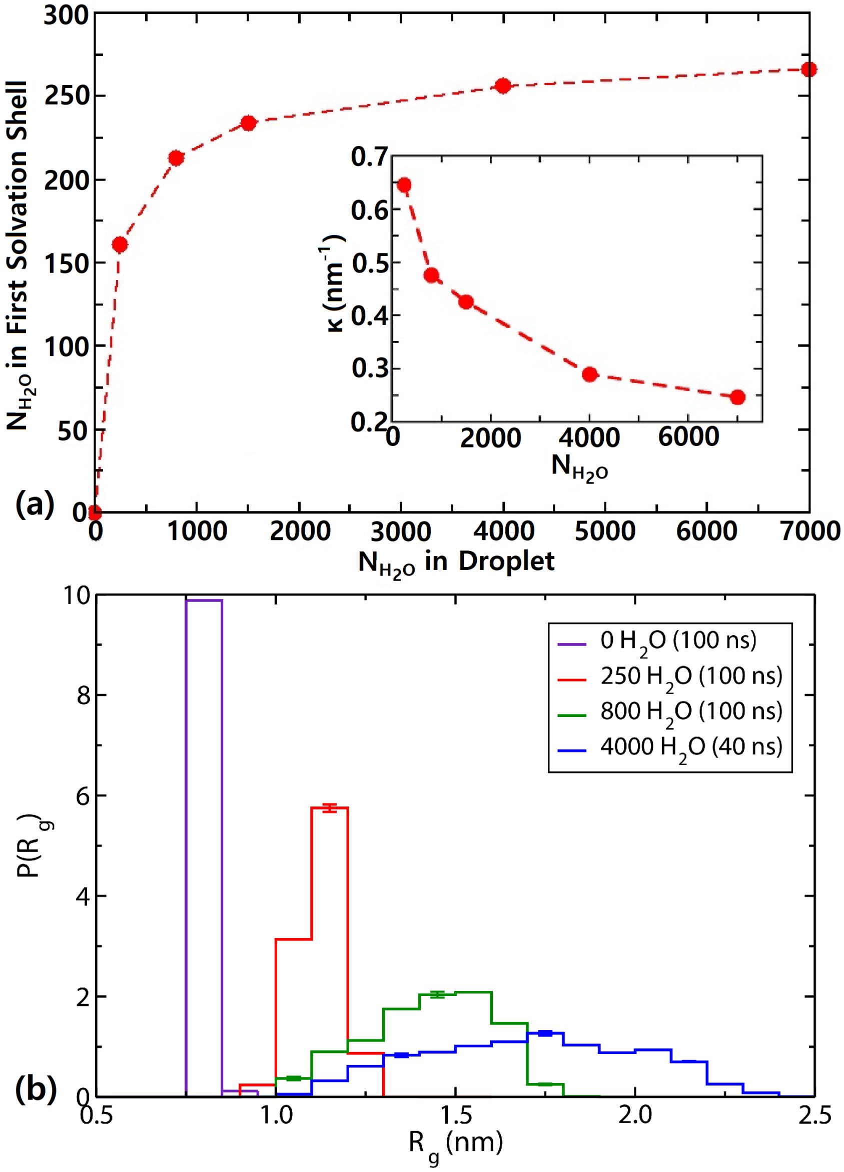


Figure A.3: (a) Number of H_2O molecules in the first solvation shell of PEG64 as a function of the droplet size. The inset shows the change in the droplet curvature (κ) *vs.* N_{H_2O}. (b) Probability density (P) of the radius of gyration (R_g) of PEG64 in aqueous droplets of various sizes. The length of the production run is indicated in the legend and the bin size is 0.1 nm.

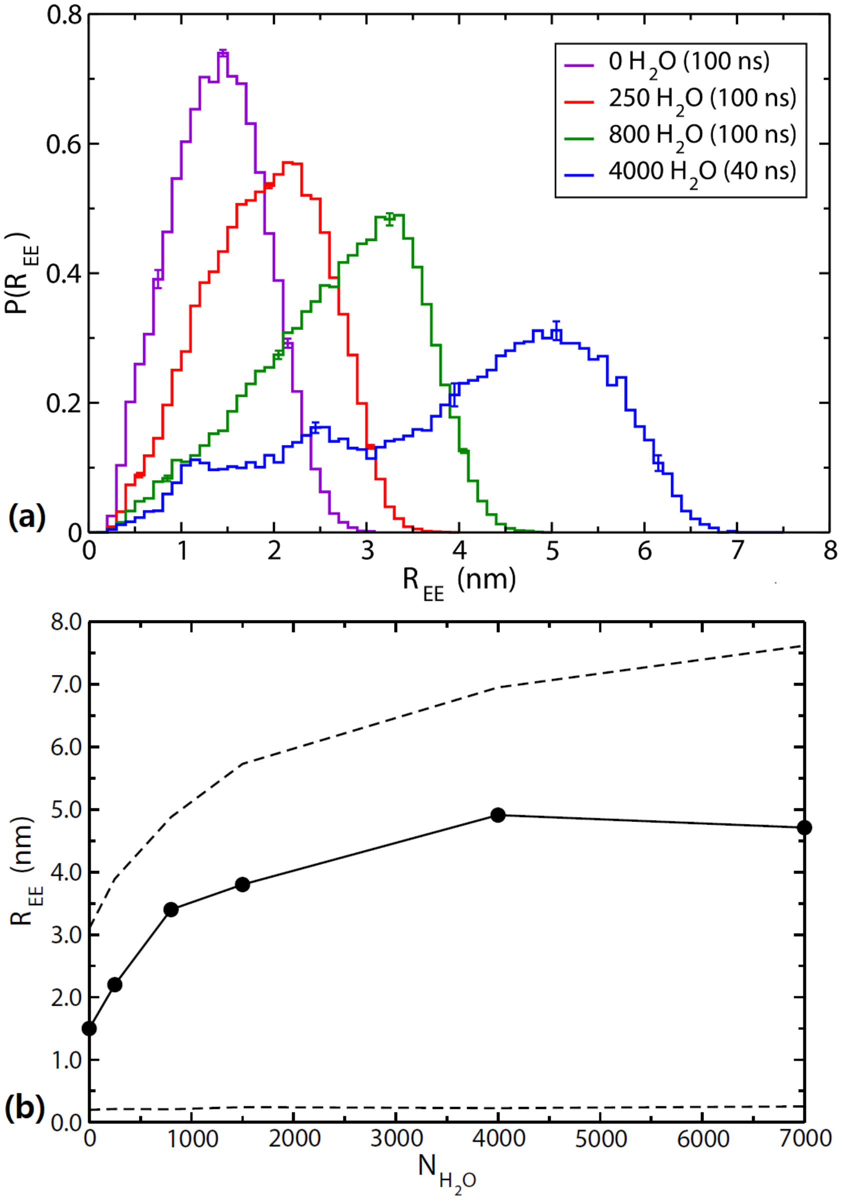


Figure A.4: (a) Probability density (P) of the end-to-end distance (R_{EE}) of PEG64 in aqueous droplets of various sizes. The length of the production run is indicated in the legend and the bin size is 0.1 nm. (b) Typical plot showing the convergence of (the most probable) R_{EE} of PEG64 as the number of water molecules in an aqueous droplet ($N_{\mathrm{H_2O}}$) increases. The dotted lines indicate the maximum (above) and minimum (below) values of R_{EE}.

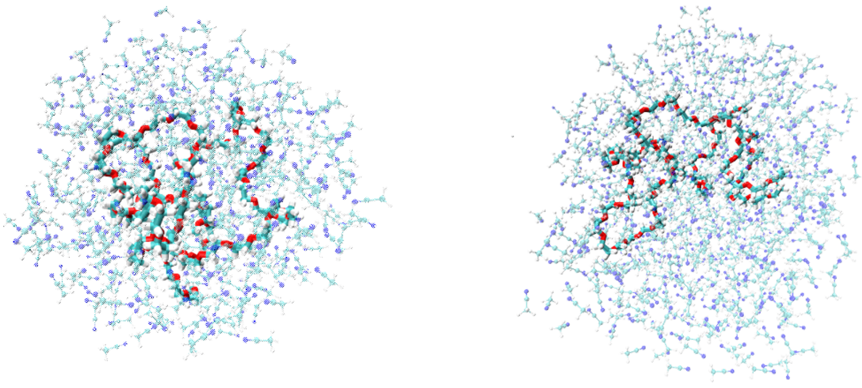
A.3 Macromolecules in the Interior of an Acetonitrile Droplet

Differently from PEG64 in aqueous droplets, it is consistently observed that the macromolecule settles in the interior of the MeCN droplets. Typical configurations of the droplets are shown in Figure A.5.

In general, the preferential location of PEG64 on the surface *vs.* in the interior of droplets of various solvents has its roots in the relative strengths between the solvent-monomer and solvent-solvent interactions. The complete solubility of PEG in water and many organic solvents such as MeCN, MeOH, benzene, and chloroform has been widely recognized experimentally [2, 3]. MeCN molecules lack the hydrogen-bonding ability, and therefore, it is dipole-dipole interactions that prevail for solvent-solvent and solvent-monomer interactions. The lack of the ability to form strong hydrogen bonds between PEG and MeCN may be compensated by the high dipole moments of the individual solvent molecules. Hakem and Lal [3] computed the solubility parameters of PEO in H_2O and MeCN solutions at room temperature by applying the Hildebrand equation, and they expected lower polymer solubility in MeOH than in MeCN. In our previous study [4], we found that PEG resides inside a MeOH droplet, regardless of the presence of charge carriers. Moreover, MeCN molecules share a common methyl group that water molecules do not possess. These two factors, *i.e.*, the similarity in the structure and the intermolecular interactions, may cause the solvation of PEG in the interior of MeCN droplets.

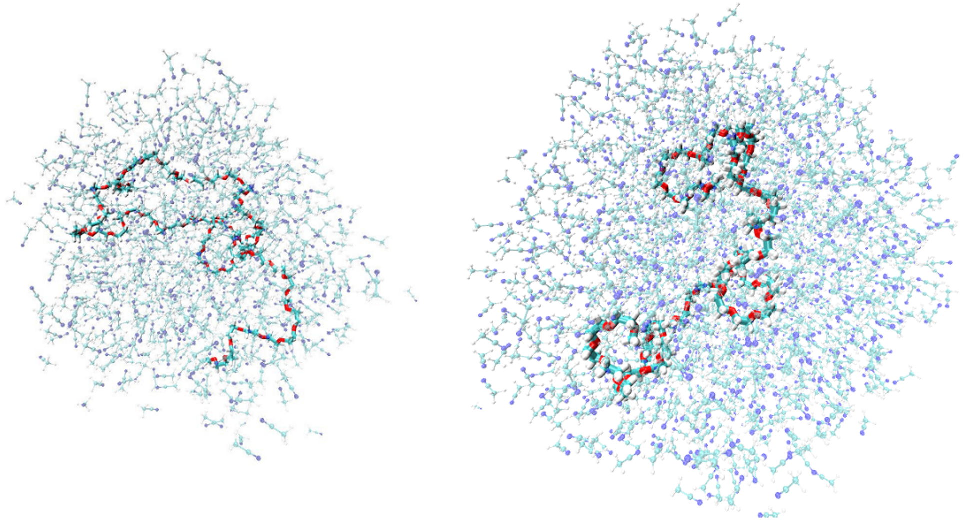
Since PEG64 lies in the interior of the droplet, we conjecture that its dimensions are affected largely by the diameter of the droplet. We centre the following discussion around MeCN: Figures A.6 (a) and (b) show the R_g and R_{EE} distributions of PEG64 in MeCN droplets as a function of the droplet size. The R_g and R_{EE} distributions reveal the strong correlation of the conformational dimensions to the size of the hosting droplet. In addition, their fluctuations also increase with the size of the droplet, implying more conformational freedom granted to the polymer simply due to more physical space available to it. In MeCN, the dimensions of PEG64 with respect to the number of solvent molecules converge faster than those observed in H_2O.

The convergence of PEG64 dimensions to an upper limit is in agreement with the findings of Cifra *et al.* [5] who demonstrated that the mean R_{EE} and R_g of a polymer in a spherical cavity



(a) PEG64 + 450 MeCN

(b) PEG64 + 800 MeCN



(c) PEG64 + 750 MeCN

(d) PEG64 + 1500 MeCN

Figure A.5: Snapshots of configurations of PEG64 in MeCN droplets of different sizes: (a) 450 MeCN, (b) 800 MeCN, (c) 750 MeCN, (d) 1500 MeCN

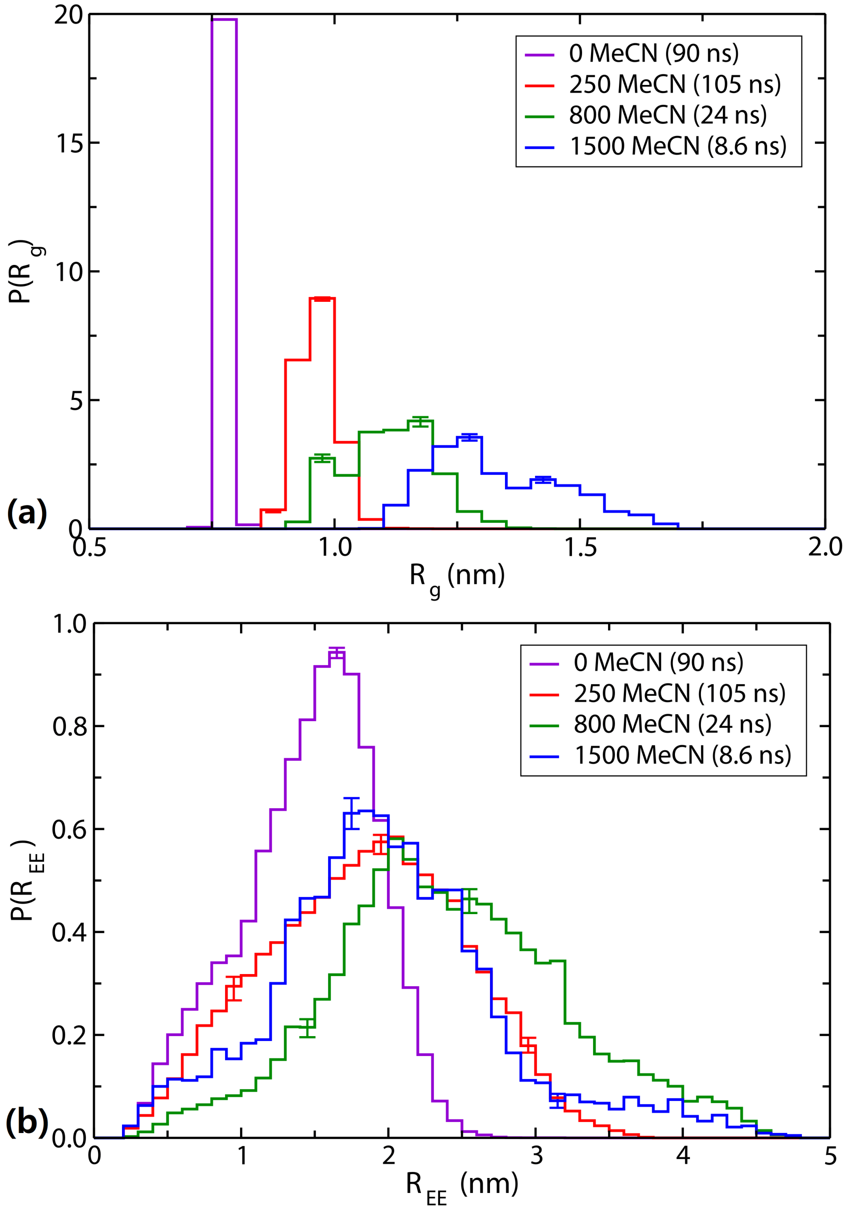


Figure A.6: Probability density (P) of (a) the radius of gyration (R_g) and (b) the end-to-end distance (R_{EE}) of PEG64 in MeCN droplets of various sizes. The length of the production run is indicated in the legend. The bin sizes for R_g and R_{EE} are 0.05 nm and 0.1 nm, respectively.

converge to an upper limit as the sphere radius increases. The same trend was observed for polymers with different stiffness parameters [5]. In those simulations a hybrid coarse-grained model of semi-flexible chains was employed, which comprises bead units and stiff springs for effective bonds. The sphere restricts the motions of the monomers in all directions when its size is quite small compared to the size of the macromolecule, yet when the radius is fairly large, it can mimic the bulk environment. The confinement effect exerted by the droplet's finite size on the dimensions of the macromolecule is a general feature that is consistent with the findings of other theoretical studies of the effect of a cavity on polymer dimensions [5–12].

In summary, the strong droplet-size-dependent conformational dimensions of PEG64 and their convergence to an upper threshold are common features in the solvents of water and acetonitrile. These findings indicate that the confinement effect is not solvent-specific, and it is completely independent of the location of the macromolecule in a droplet. The confinement effect manifests itself in the most probable values of the R_g and R_{EE} and the extent of their fluctuations. We found that the most probable R_g and R_{EE} of PEG64 in water droplets decrease by approximately 50% and 70%, respectively, as the radius of the droplet decreases from 3.8 nm to the zero solvent molecules (*i.e.*, dry PEG64 conformation). Also, in MeCN droplets, PEG64 shows a decline in the most probable R_g and R_{EE} values, approximately by 39% and 12%, respectively, as the radius of the droplet drops from 3.2 nm to zero. The transition from the extended conformation of PEG64 in larger droplets to the dry state is continuous. Because of the convergence in the most probable values of R_g and R_{EE} in droplets with radius larger than 3.3 nm and 3.2 nm for water and MeCN, respectively, we demonstrated that the droplet confinement plays a role in droplets with radius smaller than 3.2 nm. When PEG64 is close to dryness, the confinement effect diminishes.

References

[1] Prasitnok, K. & Wilson, M. R. A coarse-grained model for polyethylene glycol in bulk water and at a water/air interface. *Phys. Chem. Chem. Phys.* **15**, 17093–17104 (2013).

[2] Zhou, P. & Brown, W. Static and dynamic properties of poly(ethylene oxide) in methanol. *Macromolecules* **23**, 1131–1139 (1990).

[3] Hakem, I. F. & Lal, J. Polyelectrolyte-like behaviour of poly(ethylene-oxide) solutions with added monovalent salt. *Europhys. Lett.* **64**, 204 (2003).

[4] Soltani, S., Oh, M. & Consta, S. Effect of solvent on the charging mechanisms of poly(ethylene glycol) in droplets. *J. Chem. Phys.* **142**, 114307 (2015).

[5] Cifra, P., Benková, Z. & Bleha, T. Persistence lengths and structure factors of wormlike polymers under confinement. *J. Phys. Chem. B* **112**, 1367–1375 (2008).

[6] Shin, J., Cherstvy, A. G. & Metzler, R. Kinetics of polymer looping with macromolecular crowding: effects of volume fraction and crowder size. *Soft Matter* **11**, 472–488 (2015).

[7] Shin, J., Cherstvy, A. G. & Metzler, R. Polymer looping is controlled by macromolecular crowding, spatial confinement, and chain stiffness. *ACS Macro Lett.* **4**, 202–206 (2015).

[8] Cifra, P., Benkova, Z. & Bleha, T. Effect of confinement on properties of stiff biological macromolecules. *Faraday Discuss.* **139**, 377–392 (2008).

[9] Cifra, P. & Bleha, T. Simulation of chain organization in encapsulated polymers. *Macromol. Symp.* **296**, 336–341 (2010).

[10] Morrison, G. & Thirumalai, D. Semiflexible chains in confined spaces. *Phys. Rev. E* **79**, 011924 (2009).

[11] Cerdà, J. J., Sintes, T. & Chakrabarti, A. Excluded volume effects on polymer chains confined to spherical surfaces. *Macromolecules* **38**, 1469–1477 (2005).

[12] Spakowitz, A. J. & Wang, Z.-G. Semiflexible polymer confined to a spherical surface. *Phys. Rev. Lett.* **91**, 166102 (2003).

Appendix B

Derivation of the Rayleigh Criterion

The total energy of a charged droplet is the sum of its surface energy E_{surf} and electrostatic energy E_{Coul}:

$$E = E_{\text{surf}} + E_{\text{Coul}} = \gamma A + \frac{1}{2} Q\Phi \tag{B.1}$$

where Q, γ, A, and Φ denote the net charge, surface tension, surface area, and surface electrostatic potential, respectively, of the droplet. Under the assumption that the droplet is a conductor, the electrostatic potential is constant. The perturbations of a spherical droplet from its equilibrium geometry are expanded in terms of spherical harmonics:

$$\rho(\sigma) = R_0 + \sum_{l>0,m_l} a_{l,m_l} Y_{l,m_l}(\sigma) \tag{B.2}$$

where ρ is the distance from the centre, R_0 is the $l = 0$ term in the expansion of $\rho(\sigma)$ and $\sigma = (\phi, \theta)$ is the spherical coordinates. Note that, in general, $R_0 \neq R$ where R is the radius of an unperturbed spherical droplet. The general solution for the electrostatic potential in a vacuum decaying at infinity in spherical polar coordinates is given by the solution of

$$\Phi(r, \sigma) = \frac{Q}{4\pi\varepsilon_0 r} + \sum_{l>0,m_l} \frac{b_{l,m_l}}{r^{l+1}} Y_{l,m_l}(\sigma) \tag{B.3}$$

where $Y_{l,m_l}(\sigma)$ is the spherical harmonics and r is the distance from the centre of the conductor. Q/r in Equation C.3 is the $l = 0$ term of the expansion of $\Phi(r, \sigma)$ in terms of $Y_{l,m_l}(\sigma)$. The first term is the dominant term in the expansion. Hereafter, we use the following notation

$$\delta r = \sum_{l>0,m_l} a_{l,m_l} Y_{l,m_l}(\sigma). \tag{B.4}$$

We find the relation between the coefficients a_{l,m_l} and b_{l,m_l} using the fact that on the surface of the droplet, the electrostatic potential is constant:

$$\widetilde{\Phi}(\sigma) = \Phi|_{\rho,\sigma} = C \tag{B.5}$$

where C is a constant, and using the orthogonality condition of the spherical harmonics

$$\int_{S^2} \widetilde{\Phi}(\sigma) Y_{l,m_l}(\sigma) d\sigma = 0 \tag{B.6}$$

for all $l > 0$, where $d\sigma = \sin\theta d\theta d\phi$ is the area of the spherical surface element. In Equation C.6, we assume that the fluctuations are small and we use Taylor expansions of $1/r$ and $1/r^{l+1}$ at $r = R_0$:

$$\begin{aligned} \frac{1}{r} &= \frac{1}{R_0} - \frac{1}{R_0^2}\delta r + \frac{1}{R_0^3}(\delta r)^2 \\ \frac{1}{r^{l+1}} &= \frac{1}{R_0^{l+1}} - \frac{l+1}{R_0^l}\delta r \end{aligned} \tag{B.7}$$

Explicitly using Equation C.3 for the value of the electrostatic potential, we recast Equation C.6 as

$$\int_{S^2} \left[\frac{1}{4\pi\varepsilon_0}\frac{Q}{\rho(\sigma)} + \sum_{l'>0,m'_l} \frac{b_{l',m'_l}}{\rho(\sigma)^{l+1}} Y_{l',m'_l}(\sigma) \right] Y_{l,m_l}(\sigma) d\sigma = 0. \tag{B.8}$$

We use the Taylor series (Equation C.7) for $1/\rho(\sigma)$ and $1/\rho(\sigma)^{l+1}$. In the expansion of $1/\rho(\sigma)$, we keep terms up to the first order term in δr and the zeroth order terms in the expansion of $1/\rho(\sigma)^{l+1}$. Evaluation of the integral in Equation C.6 for every spherical harmonics with $l > 0$ yields

$$\frac{1}{4\pi\varepsilon_0}\frac{Q}{R_0^2} a_{l,m_l} = \frac{b_{l,m_l}}{R_0^{l+1}} \Rightarrow a_{l,m_l} = 4\pi\varepsilon_0 \frac{b_{l,m_l}}{QR_0^{l-1}}. \tag{B.9}$$

For $l = 0$, since $\widetilde{\Phi}(\sigma)$ is a constant, we establish the following identity

$$\widetilde{\Phi}(\sigma) = \frac{1}{4\pi}\int_{S^2}\widetilde{\Phi}(\sigma)d\sigma. \tag{B.10}$$

Using higher order terms in the Taylor series (Equation C.7), Equations C.3, B.9, and B.10 give the following relation:

$$\widetilde{\Phi}(\sigma) = \frac{1}{4\pi\varepsilon_0}\left[\frac{Q}{R_0} + \frac{1}{4\pi}\frac{Q}{R_0^3}\sum_{l>0,m_l} a_{l,m_l}^2 - \frac{1}{4\pi}\frac{Q}{R_0^3}\sum_{l>0,m_l}(l+1)a_{l,m_l}^2\right]. \tag{B.11}$$

Quantities R_0 and a_{l,m_l} are not independent but connected via the constant volume constraint. By using the expansion $\rho^3 = (R_0 + \delta r)^3$,

$$\frac{1}{3}\rho^3 = \frac{1}{3}R_0^3 + R_0^2\delta r + R_0(\delta r)^2 \tag{B.12}$$

and by evaluating the integral of Equation B.12 over the unit sphere, the total droplet volume can be written as

$$\int_{S^2}\frac{1}{3}\rho^3 d\sigma = \frac{4}{3}\pi R_0^3 + R_0\sum_{l>0,m_l} a_{l,m_l}^2 = \frac{4}{3}\pi R^3. \tag{B.13}$$

Using the approximations $R^3 - R_0^3 \approx 3(R - R_0)R^2$ and $R_0 \approx R$, we reduce Equation B.13 to the following identity:

$$4\pi(R_0 - R)R \approx -\sum_{l>0,m_l} a_{l,m_l}^2. \tag{B.14}$$

Using the expansion for the first term of the expression for the potential energy (Equation C.3)

$$\frac{Q}{R_0} \approx \frac{Q}{R}\left[1 - \frac{R_0 - R}{R}\right] \tag{B.15}$$

and Equation B.14 for $R_0 - R$, we arrive at the final expression for the value of the electrostatic potential at the droplet surface:

$$\Phi_0 = \frac{1}{4\pi\varepsilon_0}\left[\frac{Q}{R} - \frac{Q}{4\pi R_0^3}\sum_{l>0,m_l}(l-1)a_{l,m_l}^2\right] \tag{B.16}$$

and the electrostatic energy is given by

$$E_{\text{Coul}} = -\frac{1}{2}Q\Phi_0 \tag{B.17}$$

Now we look into the surface term. We consider the position on the surface in the Cartesian coordinates

$$\boldsymbol{\rho} = [x\ y\ z]^T = \begin{bmatrix} \rho(\theta,\phi)\sin\theta\cos\phi \\ \rho(\theta,\phi)\sin\theta\sin\phi \\ \rho(\theta,\phi)\cos\theta \end{bmatrix}. \tag{B.18}$$

An infinitesimal area on the surface spanned by $(d\phi, d\theta)$ is defined as $dA = \left|\Delta\boldsymbol{u} \times \Delta\boldsymbol{v}\right|$ where $\Delta\boldsymbol{u} = (\partial\boldsymbol{\rho}/\partial\phi)\,d\phi$ and $\Delta\boldsymbol{v} = (\partial\boldsymbol{\rho}/\partial\theta)\,d\theta$. Then, we arrive at the following expression for the area of the surface element:

$$dA = \rho^2 \sin\theta d\theta d\phi \sqrt{1 + \frac{1}{\rho^2}\left[\left(\frac{\partial\rho}{\partial\theta}\right)^2 + \frac{1}{\sin^2\theta}\left(\frac{\partial\rho}{\partial\phi}\right)^2\right]} \tag{B.19}$$

The second term under the square root is small relative to unity for small deviations from the spherical shape. The Taylor expansion of the square root function yields the approximate expression for the elementary surface:

$$dA \approx \rho^2 d\sigma\left[1 + \frac{1}{2\rho^2}\left\{\left(\frac{\partial\rho}{\partial\theta}\right)^2 + \frac{1}{\sin^2\theta}\left(\frac{\partial\rho}{\partial\phi}\right)^2\right\}\right] \tag{B.20}$$

Finally,

$$A = \int_{S^2} dA = \int_{S^2} d\sigma\left[\rho^2 + \frac{1}{2}\left\{\left(\frac{\partial\rho}{\partial\theta}\right)^2 + \frac{1}{\sin^2\theta}\left(\frac{\partial\rho}{\partial\phi}\right)^2\right\}\right]. \tag{B.21}$$

Integration by parts leads to the following equation:

$$A = \int_{S^2} d\sigma\left(\rho^2 + \frac{1}{2}\rho\hat{L}\rho\right) \tag{B.22}$$

where $\hat{L}$ is the angular momentum operator. Using Equation C.2, we deduce that Equation B.22 yields the following:

$$A = 4\pi R_0^2 + \sum_{l>0,m_l} a_{l,m_l}^2 + \frac{1}{2}\sum_{l>0,m_l} l(l+1)a_{l,m_l}^2. \tag{B.23}$$

Using $4\pi R_0^2 = 4\pi R^2 + 8\pi R(R_0 - R)$ and Equation B.14, we further obtain

$$A = 4\pi R^2 + \frac{1}{2}\sum_{l>0,m_l}(l-1)(l+2)a_{l,m_l}^2. \tag{B.24}$$

Now the total energy is given by

$$E = \frac{1}{2}\frac{1}{4\pi\varepsilon_0}\left[\frac{Q^2}{R} - \frac{Q^2}{4\pi R^3}\sum_{l>0,m_l}(l-1)a_{l,m_l}^2\right] + \gamma\left[4\pi R^2 + \frac{1}{2}\sum_{l>0,m_l}(l-1)(l+2)a_{l,m_l}^2\right] \tag{B.25}$$

where γ is the surface tension of the system. Thus, δE is given by

$$\delta E = \frac{1}{2}\sum_{l>0,m_l}(l-1)\left[(l+2)\gamma - \frac{Q^2}{(4\pi)^2\varepsilon_0 R^3}\right]a_{l,m_l}^2 \tag{B.26}$$

Note the absence of the $l = 1$ term. The $l = 1$ term vanishes because the $l = 1$ mode corresponds to uniform translation of the surface. Clearly, a uniform translation of the droplet does not change the energy of the system. The first term that may allow unbounded fluctuations is the term with $l = 2$. Equating this term to zero leads to Rayleigh criterion. The droplet is stable with respect to small perturbations when

$$Q^2 < 64\pi^2\varepsilon_0\gamma R^3. \tag{B.27}$$

Appendix C

Molecular Dynamics Methods

C.1 Molecular Dynamics

The motion of an object under the regime of classical mechanics is described by Newton's three laws of motion, where its time evolution can be expressed as a set of mathematical functions of dynamical variables such as spatial coordinates and time. For a system comprising multiple particles without the influence of an external force, the motion of each particle is governed by the mutual interactions between every pair of the particles, and this may be expressed explicitly as follows:

$$\boldsymbol{F}_i(t) = m_i \ddot{\boldsymbol{r}}_i(t) \tag{C.1}$$

where $\boldsymbol{F}_i(t)$ is the net force acting on particle i as a function of time t, m_i is the mass of particle i, and $\ddot{\boldsymbol{r}}_i$ is the second derivative of the position with respect to time, namely, the acceleration of particle i.

The deterministic nature of classical mechanics is implied in Equation C.1. Assuming linear motion, Equation C.1 may be solved in terms of time, leading to the most critical equation for direct molecular dynamics simulations:

$$\boldsymbol{r}_i(t) = \frac{\boldsymbol{F}_i}{m_i} t^2 + \dot{\boldsymbol{r}}_i(0)t + \boldsymbol{r}_i(0) \tag{C.2}$$

where $\dot{\boldsymbol{r}}_i(0)$ and $\boldsymbol{r}_i(0)$ represent the velocity and position of particle i at $t = 0$. If the system contains N particles, then one needs to solve sN second-order differential equations where s is the dimensionality of a space. It is deterministic as one can determine the trajectory of the system, given the initial conditions $\dot{\boldsymbol{r}}_i(0)$ and $\boldsymbol{r}_i(0)$. In order to specify the classical state of a physical system at time t, therefore, a complete set of initial positions and velocities (or momenta $\boldsymbol{p}_i(0)$) for all the particles are required in addition to the net force acting on each particle.

When a force is dependent only on the spatial position and has a corresponding potential energy, it is said to be conservative. The work done by a conservative force in moving a particle between two points is independent of the path taken by the system. If $\boldsymbol{F}_i$ is conservative, then it is equal to the negative gradient of the potential energy:

$$\boldsymbol{F}_i(t) = -\nabla_i U\left(\{\boldsymbol{r}(t)\}\right) \tag{C.3}$$

where U is the total potential energy of the many-particle system, and $\{\boldsymbol{r}\} = (\boldsymbol{r}_1, \boldsymbol{r}_2, \cdots, \boldsymbol{r}_N)$ is the set of the spatial coordinates of all the particles in the system. Assuming pairwise additivity, the total potential energy of the system can be obtained by summing over all distinct pairwise interactions:

$$U\left(\{\boldsymbol{r}\}\right) = \sum_{i}^{N-1} \sum_{j>i}^{N} u_{ij}\left(\boldsymbol{r}_{ij}\right) \tag{C.4}$$

where $u_{ij}\left(\boldsymbol{r}_{ij}\right)$ is the potential energy of a pair interaction as a function of the displacement between two particles i and j.

For soft spherical particles with closed electronic shells and no net electric charges, such as rare gas atoms, they interact with one another only via van der Waals interactions, and their pairwise potential energy is approximated with the Lennard-Jones (LJ) 12-6 potential which is given by

$$u_{ij}\left(\boldsymbol{r}_{ij}\right) = 48\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{\boldsymbol{r}_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{\boldsymbol{r}_{ij}}\right)^{6}\right] \tag{C.5}$$

where σ denotes the distance of separation between two particles i and j for which the pair potential is zero, and ε represents the strength (minimum) of the potential energy, that is, the potential well depth. These values are experimentally determined. The first term $\left(\boldsymbol{r}_{ij}\right)^{-12}$ in the square bracket describes the Pauli repulsion, predominating at very short distance due to electronic orbital overlapping, and the second term $\left(\boldsymbol{r}_{ij}\right)^{-6}$ reflects the dispersive forces or attractive interactions due to instantaneous dipoles arising from the fluctuations in electron clouds. When the pairwise potential energy is described by the LJ potential, the force acting on one particle i by interacting with the other particle j can be expressed in the following form:

$$\boldsymbol{F}_{ij} = -\frac{\partial u_{ij}\left(\boldsymbol{r}_{ij}\right)}{\partial \boldsymbol{r}_{ij}} = 24\varepsilon\left[2\left(\frac{\sigma_{ij}}{\boldsymbol{r}_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{\boldsymbol{r}_{ij}}\right)^{6}\right]\frac{\boldsymbol{r}_{ij}}{r_{ij}^{2}} = -\boldsymbol{F}_{ij} \tag{C.6}$$

The force acting on particle i by particle j is the same as the force acting on particle j by particle i in magnitude but opposite in direction according to the law of action and reaction.

In more realistic models of intermolecular interactions, the force on each particle changes as its spatial position changes or other particles with which it interacts change their positions, varying the potential energy continuously with distance. Under the influence of a continuous potential, the motions of all the particles are coupled. Due to the complexity of the potential energy of a many-body system, in general, there exists no analytical method to integrate the differential form of the equations of motion. Molecular dynamics simulations, therefore, must employ an iterative numerical algorithm, called a finite-difference method, to approximate the solution and thus generate molecular dynamics trajectories. In this method, the total potential energy is assumed to remain constant in each time frame.

The time average of a physical observable B over a numerical trajectory generated by molecular dynamics simulations can be expressed as the following mathematical form:

$$\overline{B} = \frac{1}{\tau_s}\int_0^{\tau_s} B\left(\boldsymbol{r}^N(t), \boldsymbol{p}^N(t)\right)dt = \frac{1}{S}\sum_{i=1}^{S} B\left(\boldsymbol{r}^N(t), \boldsymbol{p}^N(t)\right) \tag{C.7}$$

where τ_r is the length of time elapsed, S is the total number of time steps, and $\boldsymbol{p}$ is linear momenta.

The primary objective of molecular dynamics simulations is to generate molecular trajectories over a finite time. For any arrangement of particles in the system, the net force acting on each particle due to interatomic interactions with other particles may be calculated by differentiating the total potential energy function. From the net force on each atom, its acceleration is determined via Newton's second law. The numerical integration of the equations of motion yields a trajectory that describes the time evolution of the system, namely, how the positions, velocities, and acceleration of the particles vary with time, and from which the time average values of physical quantities can be calculated using Equation C.7.

C.2 Numerical Integrators

A finite-difference method is a numerical algorithm used to integrate the equations of motion and thus advance molecular trajectories with continuous potential models over a small fixed time increment:

$$\boldsymbol{r}(t_0) \rightarrow \boldsymbol{r}(t_0 + \delta t) \rightarrow \boldsymbol{r}(t_0 + 2\delta t) \rightarrow \cdots \rightarrow \boldsymbol{r}(t_0 + S\,\delta t)$$

where t_0 is the initial time at which the numerical integration begins, δt is the length of each time step, and S is the total number of time steps. A wide variety of numerical integrators have been developed and implemented into molecular dynamics simulations to generate numerical trajectories of particles. All the algorithms assume that, since molecular trajectories are smooth and continuous, dynamic quantities such as position, velocity, and acceleration can be approximated as Taylor series expansions:

$$\boldsymbol{r}(t + \delta t) \approx \boldsymbol{r}(t) + \boldsymbol{v}(t)\delta t + \frac{1}{2}\boldsymbol{a}(t)\delta t^2 \tag{C.8}$$

$$\boldsymbol{v}(t + \delta t) \approx \boldsymbol{v}(t) + \boldsymbol{a}(t)\delta t + \frac{1}{2}\boldsymbol{b}(t)\delta t^2 \tag{C.9}$$

$$\boldsymbol{a}(t + \delta t) \approx \boldsymbol{a}(t) + \boldsymbol{b}(t)\delta t + \frac{1}{2}\boldsymbol{c}(t)\delta t^2 \tag{C.10}$$

where $\boldsymbol{b}(t)$ and $\boldsymbol{c}(t)$ are third and fourth derivatives of the position vector $\boldsymbol{r}(t)$ with respect to time, respectively. For the duration of each time step, the force acting on each particle is calculated from the potential energy and assumed to be constant. Once the acceleration of each particle is calculated according to Newton's second law, its position and velocity at a next time step are predicted by combining it with the position and velocity at the current time.

The Verlet algorithm, one of the most commonly used methods, employs the position and acceleration at the current time step and the position from the previous time step to calculate a new position:

$$\boldsymbol{r}\,(t+\delta t) = 2\boldsymbol{r}(t) - \boldsymbol{r}\,(t-\delta t) + \boldsymbol{a}(t)\delta t^2 \tag{C.11}$$

which is obtained by rearranging the resultant equation after adding the following Taylor expansions:

$$\boldsymbol{r}\,(t+\delta t) \approx \boldsymbol{r}(t) + \boldsymbol{v}(t)\delta t + \frac{1}{2}\boldsymbol{a}(t)\delta t^2 \tag{C.12}$$

$$\boldsymbol{r}\,(t-\delta t) \approx \boldsymbol{r}(t) - \boldsymbol{v}(t)\delta t + \frac{1}{2}\boldsymbol{a}(t)\delta t^2 \tag{C.13}$$

At the initialization stage of molecular dynamics simulations (that is, $t = 0$), it is necessary to employ some means to obtain positions at $t = -\delta t$. One possible way is to use a truncated Taylor expansion, namely,

$$\boldsymbol{r}(-\delta t) \approx \boldsymbol{r}(0) - \boldsymbol{v}(0)\delta t \tag{C.14}$$

where the initial position and velocity of each particle may be assigned according to the Boltzmann distribution. Since the additional computation of $\boldsymbol{r}(-\delta t)$ must precede the initialization stage, the Verlet algorithm is not self-starting. Due to the lack of an explicit velocity term in the equation, the velocity is computed after the new position is calculated at the next time step:

$$\boldsymbol{v}(t) = \frac{\boldsymbol{r}\,(t+\delta t) - \boldsymbol{r}\,(t-\delta t)}{2\delta t} \tag{C.15}$$

The leap-frog algorithm, one of many variants of the Verlet algorithm, uses the following equations:

$$\boldsymbol{r}\,(t+\delta t) = \boldsymbol{r}(t) + \boldsymbol{v}\left(t + \frac{1}{2}\delta t\right)\delta t \tag{C.16}$$

$$\boldsymbol{v}\left(t + \frac{1}{2}\delta t\right) = \boldsymbol{v}\left(t - \frac{1}{2}\delta t\right) + \boldsymbol{a}(t)\delta t \tag{C.17}$$

and the velocity at time t is obtained from

$$\boldsymbol{v}(t) = \frac{\boldsymbol{v}\left(t + \frac{1}{2}\delta t\right) + \boldsymbol{v}\left(t - \frac{1}{2}\delta t\right)}{2} \tag{C.18}$$

Since the velocity at $t + 0.5\delta t$ is computed first to find the position at $t + \delta t$, the velocity calculation jumps over the position computation which leaps over the next velocity calculation as the name ‘leap-frog’ implies. Since position and velocity are not synchronized when they are defined, it is impossible to calculate the kinetic energy at the same time as the potential energy contribution to the total energy of the system.

Appendix D

C++ Codes for Monte Carlo Simulations

D.1 Simple Lennard-Jones System

Particles in a simple Lennard-Jones system (such as clusters of noble gas atoms) communicate with one another only via van der Waals interactions. The van der Waals interactions are often modeled by the so-called Lennard-Jones 12-6 potential. The Lennard-Jones parameters for each atomic site was taken from those of krypton (Kr). Two different temperatures were used: 50 K and 100 K. The total number of steps is 500000. For each temperature, the radial distribution function (RDF) was computed (Figure D.1). The RDF includes information about the structure of a solid or a liquid (condensed phase).

In Figure D.1, the peaks indicate the location of the shells of neighbouring particles, and the area under each peak gives the number of particles in the corresponding shell. It is clearly seen that, in both cases, the first peak is located at r = 0.39 nm which is close to the atomic radius of Kr. The peaks become smaller as the distance from the central atom is larger, since there is less "communication" (or correlation in the behaviour) between the atom to those located far in distance (*e.g.*, they do not exhibit any structural periodicity as in a typical solid lattice). The curves are all dying off to zero regardless of temperature due to the presence of the physical boundary (*i.e.*, they are clusters). As expected, the RDF is strongly dependent on temperature; the higher temperature, the smaller number and intensity of the peaks (Figure D.1 (b)).

The C++ code for the Monte Carlo simulations was written as follows:

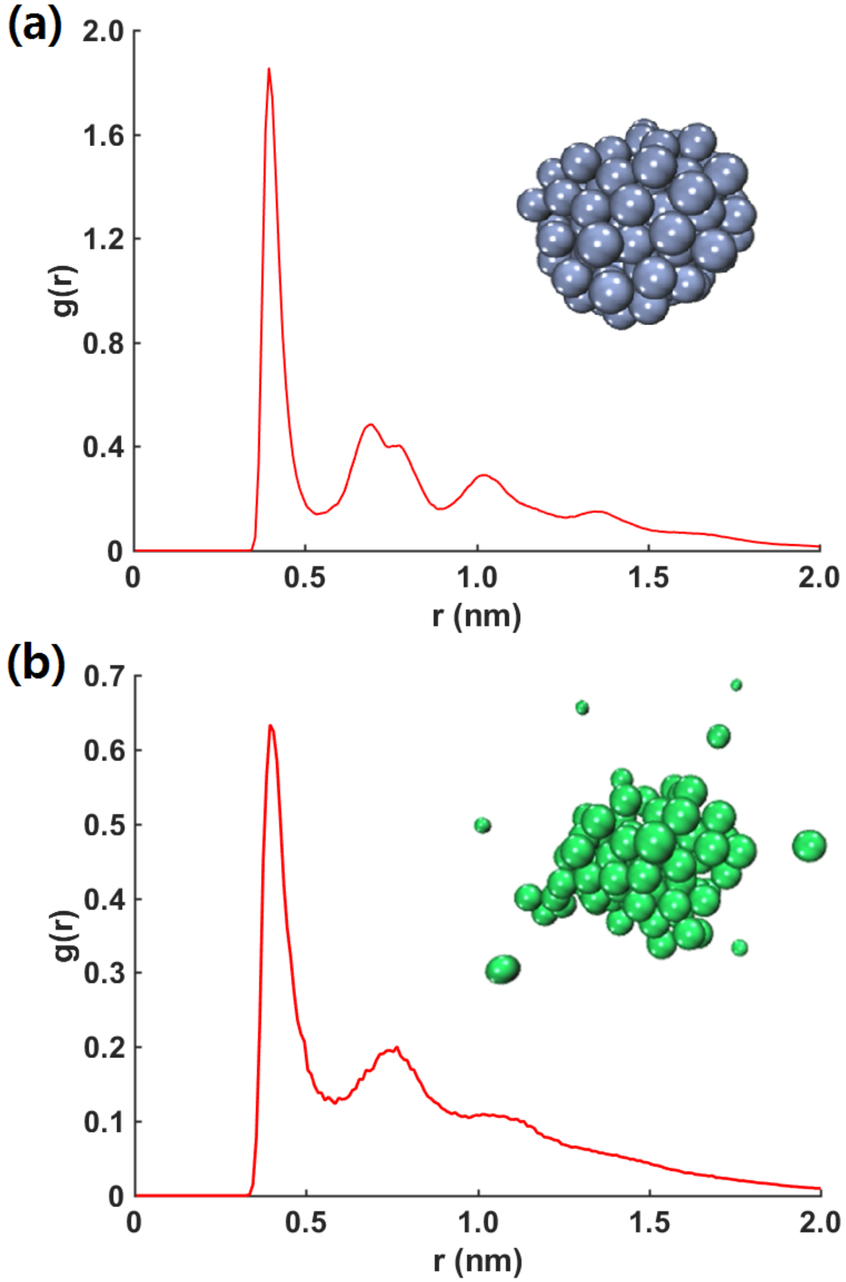


Figure D.1: The radial distribution functions of krypton clusters (composed of 125 Kr atoms) at (a) 50 K and (b) 100 K. The snapshots of the clusters (blue particles for Kr at 50 K and green particles for Kr at 100 K) are embedded.

```
#include ⟨iostream⟩
#include ⟨cstdlib⟩
#include ⟨cmath⟩
#include ⟨fstream⟩
#include ⟨random⟩
#include ⟨iomanip⟩

using namespace std;

// Constants

const int number_of_particles = 125;
const int lattice_size = 5;
const double lattice_distance = 6.0;
const double max_translation = 1.5;
const double temperature_kelvin = 100.0;
const double sigma = 3.624; // 3.624 angstroms
const double epsilon = 0.317; // 0.317 kcal mol-1
const double k_boltz = 1.98720650419155E-3; // kcal mol-1 K-1
const int total_moves = 500000;

ofstream fileout0(”b_Trajectory_Krypton_100K.xyz”);
ofstream fileout1(”b_Conditions_Results_100K.txt”);
ofstream fileout2(”b_Energy_Distribution_100K.rtf”);

// Structures

struct position
{
    double x;
    double y;
    double z;
};

position krypton[150];
position temporary_position;

// Function Declarations

void record_conditions_results();
void position_initialization();
void random_translation_scheme();
double energy_calculation();
int random_number_generator_atom();
double random_number_generator_motion();
```

```
double random_number_generator_scheme();

int main()
{
    record_conditions_results();
    position_initialization();
    random_translation_scheme();

    for (int i = 0; i < number_of_particles; i++)
    {
        cout << ”Kr \t” << krypton[i].x << ”\t” << krypton[i].y << ”\t” << krypton[i].z
<< endl;
    }

    return 0;
}

void record_conditions_results()
{
    fileout1 << ”—————(CONDITIONS ONLY)—————” << endl;
    fileout1 << ”Number of Kr Atoms: ” << number_of_particles << endl;
    fileout1 << ”Lattice Size: ” << lattice_size << endl;
    fileout1 << ”Lattice Distance (A): ” << lattice_distance << endl;
    fileout1 << ”Maximum Translation (A): ” << max_translation << endl;
    fileout1 << ”Temperature (K): ” << temperature_kelvin << endl;
    fileout1 << ”Sigma (A): ” << sigma << endl;
    fileout1 << ”Epsilon (kcal/mol): ” << epsilon << endl;
    fileout1 << ”Boltzmann Constant (kcal/mol.K): ” << k_boltz << endl;
    fileout1 << ”Total Moves: ” << total_moves << endl;
    fileout1 << ”——————(COMPLETE!)——————” << endl;
}

void position_initialization()
{
    fileout0 << number_of_particles << endl;
    fileout0 << ”\n”;

    int m = 0;

    temporary_position.x = 0;
    for (int i = 0; i < lattice_size; i++)
    {
        temporary_position.x + = lattice_distance;
        temporary_position.y = 0;
        for (int j = 0; j < lattice_size; j++)
```

```
            {
                temporary_position.y + = lattice_distance;
                temporary_position.z = 0;
                for (int k = 0; k < lattice_size; k++)
                {
                    temporary_position.z + = lattice_distance;

                    krypton[m].x = temporary_position.x;
                    krypton[m].y = temporary_position.y;
                    krypton[m].z = temporary_position.z;

                    m++;
                }
            }
        }
        for (int i = 0; i < number_of_particles; i++)
        {
            fileout0 << ”Kr \t” << krypton[i].x << ”\t” << krypton[i].y << ”\t” << krypton[i].z << endl;
        }
    }

    void random_translation_scheme()
    {
        int n_accept = 0;
        int n_reject = 0;

        for (int move = 1; move <= total_moves; move++)
        {
            // Pick a random atom
            int m = random_number_generator_atom();

            // Calculate old energy
            double old_energy = energy_calculation();

            // Save old position
            double old_information[4] = { krypton[m].x, krypton[m].y, krypton[m].z, old_energy };

            // Make the move
            double delta_x = random_number_generator_motion();
            double delta_y = random_number_generator_motion();
            double delta_z = random_number_generator_motion();

            krypton[m].x = krypton[m].x + delta_x;
```

```
            krypton[m].y = krypton[m].y + delta_y;
            krypton[m].z = krypton[m].z + delta_z;

            cout << ”Atom ” << m << ” Chosen: dx = ” << delta_x << ”\t dy = ” << delta_y
<< ”\t dz = ” << delta_z << endl;

            // Calculate new energy
            double new_energy = energy_calculation();

            cout << ”Old Energy: ” << old_energy << ”\t” << ”New Energy: ” << new_energy
<< endl;
            cout << ”\n”;

            bool accept = false;

            // Automatically accept when energy goes down
            if (new_energy <= old_energy)
            {
                accept = true;
                n_accept + = 1;
                fileout0 << number_of_particles << endl;
                fileout0 << ”\n”;
                for (int i = 0; i < number_of_particles; i++)
                {
                    fileout0 << ”Kr \t” << krypton[i].x << ”\t” << krypton[i].y << ”\t” <<
krypton[i].z << endl;
                }
                fileout2 << move << ”\t” << new_energy << endl;
            }
            else
            {
                // Test exp(-dE/kT) >= rand[0,1]
                double p = exp((old_energy - new_energy) / (k_boltz*temperature_kelvin));
                double q = random_number_generator_scheme();
                if (p >= q)
                {
                    accept = true;
                    n_accept + = 1;
                    fileout0 << number_of_particles << endl;
                    fileout0 << ”\n”;
                    for (int i = 0; i < number_of_particles; i++)
                    {
                        fileout0 << ”Kr \t” << krypton[i].x << ”\t” << krypton[i].y << ”\t” <<
krypton[i].z << endl;
                    }
```

```
                fileout2 << move << ”\t” << new_energy << endl;
            }
            else
            {
                accept = false;
                n_reject + = 1;
                krypton[m].x = old_information[0];
                krypton[m].y = old_information[1];
                krypton[m].z = old_information[2];
            }
            if ((move % 1000) == 0)
            {
                fileout1 << ”Step ” << move << ”\t” << ”Energy (kcal/mol): ” << setprecision(2) << fixed << old_energy
                << ”\t% Acceptances: ” << double(n_accept * 100 / move)
                << ”\t% Rejections: ” << double(n_reject * 100 / move) << endl;
            }
        }
    }
    cout << ”Total Number of Acceptances: ” << n_accept << endl;
    cout << ”Total Number of Rejections: ” << n_reject << endl;
    cout << ”\n”;
}

// LJ Energy = 4*epsilon*[(sigma/r)ˆ12 - (sigma/r)ˆ6]

double energy_calculation()
{
    double delta_x, delta_y, delta_z;
    double interatomic_length_square, interatomic_energy;
    double sigma2_divide_r2;
    double sigma_square = sigma*sigma;
    double total_potential_energy = 0;

    for (int i = 0; i < number_of_particles−1; i + +)
    {
        for (int j = i + 1; j < number_of_particles; j + +)
        {
            delta_x = krypton[i].x - krypton[j].x;
            delta_y = krypton[i].y - krypton[j].y;
            delta_z = krypton[i].z - krypton[j].z;

            interatomic_length_square = delta_x*delta_x + delta_y*delta_y + delta_z*delta_z;
            sigma2_divide_r2 = sigma_square / interatomic_length_square;
```

```
                interatomic_energy = 4 * epsilon * (pow(sigma2_divide_r2, 6) - pow(sigma2_divide_r2, 3));

                total_potential_energy += interatomic_energy;
            }
        }
        return total_potential_energy;
    }

    int random_number_generator_atom()
    {
        uniform_real_distribution⟨double⟩ dist(0, number_of_particles + 0.5); // Type of random number distribution
        mt19937 rng; // Mersenne Twister: Good quality random number generator
        rng.seed(random_device{}()); // Initialize with non-deterministic seeds

        int atom = int(dist(rng));

        return atom;
    }

    double random_number_generator_motion()
    {
        double max = max_translation;
        double min = -max_translation;

        uniform_real_distribution⟨double⟩ dist(min, max); // Type of random number distribution
        mt19937 rng; // Mersenne Twister: Good quality random number generator
        rng.seed(random_device{}()); // Initialize with non-deterministic seeds

        double motion = dist(rng);

        return motion;
    }

    double random_number_generator_scheme()
    {
        uniform_real_distribution⟨double⟩ dist(0, 1); // Type of random number distribution
        mt19937 rng; // Mersenne Twister: Good quality random number generator
        rng.seed(random_device{}()); // Initialize with non-deterministic seeds

        double scheme = dist(rng);

        return scheme;
    }
```

D.2 Specifically Interacting Particles Under Spherical Confinement

A system composed of ten particles, only two of which (red and blue particles in Figure D.2) interact specifically with each other, is confined in a spherical cavity of radius of 8 Å. For the purpose of comparison, a system consisting of the two specifically interacting particles in the absence of other crowding particles is also prepared. All particles interact with one another in a nonspecific manner, and the other eight particles (purple) behave as crowders in the confining environment. The specific interaction was modeled by a square-well potential with varying well depths (ϵ_s) but a constant well width (0.5 Å). A Lennard-Jones potential was used to model the nonspecific interactions as for the argon and krypton clusters. It is interesting to study the effect of the strength of the specific interaction (*i.e.*, the depth of the square-well potential) on the binding affinity between the two specifically interacting particles. For this study, the total number of trial moves was 1000000, and the maximum distance for a translational motion of the selected particle was set to 1.5 Å. The temperature was 130 K. The ”radius” of each particle and the well depth of the Lennard-Jones potential were taken from those of krypton (*i.e.*, 3.624 Å and 0.317 kcal/mol, respectively). The depth of the square-well potential varied from 0 to −3 kcal/mol. The frequency of the binding of the specifically interacting particles f_{bound} was computed as the ratio of the number of configurations where the particles exist as a bound complex to the total number of accepted trial moves.

In Figure D.2), it is clearly seen that the increase in ϵ_s lead to the increase in f_{bound}, showing higher binding affinity between the two specifically interacting particles. Interestingly, the increase is not in a linear manner; the rate of the increase is small, reaching a plateau in the regime of very small and large ϵ_s values (*i.e.*, $\epsilon_s \leq 0.5$, $\epsilon_s \geq 2.0$ for the blue dots, $\epsilon_s \leq 0.2$, $\epsilon_s \geq 2.0$ for the red stars), while very sharp in the regime of intermediate ϵ_s values. This trend may be commonly observed regardless of the presence of the crowders. However, the effect of the presence of the crowders becomes more pronounced by comparing the two systems; crowding enhances the binding affinity of the two particles as shown in larger f_{bound} values in the presence of the nonspecific interactions by the crowders. And the crowding effect is weak or disappears when ϵ_s is relatively small or large. These findings are critical in understanding chemical

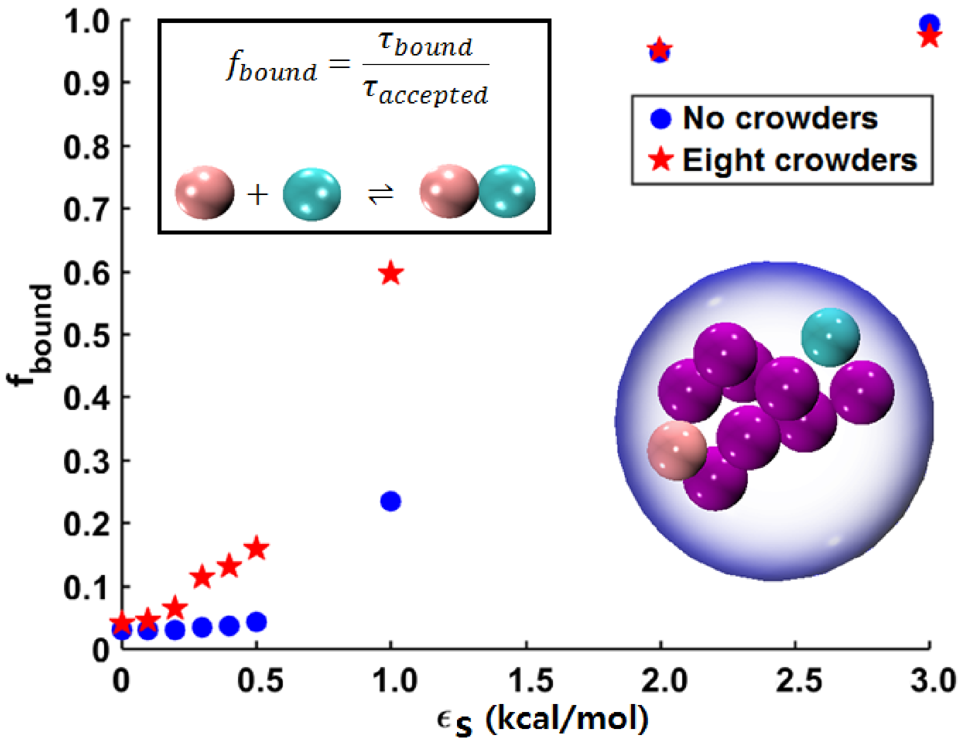


Figure D.2: Plot showing the frequency of the bound state f_{bound} as a function of ϵ_s. Blue dots correspond to the system in the absence of the crowders, and the red stars to the system containing the eight crowders. The upper inset shows the simple equation used to compute the frequency of the bound state and a schematic picture of the equilibrium between the bound and unbound states. The lower inset is a typical snapshot of the system of the particles in a special cavity.

reactions under confinement, electrospray ionization, and biological crowding environments such as vesicles, liposomes, and cells.

The C++ code written for the Monte Carlo simulations is presented in the following pages.

```
#include ⟨iostream⟩
#include ⟨cstdlib⟩
#include ⟨cmath⟩
#include ⟨fstream⟩
#include ⟨random⟩
#include ⟨limits⟩
#include ⟨iomanip⟩

using namespace std;

// Artificial = Ar, Brtificial = Br, Nonspecific = Kr
// Constants

const int number_of_nonspecific_particles = 8;
const int lattice_size = 2;
const double lattice_distance = 6.0;
const double cavity_radius = 8; // Spherical boundary condition
const double well_width = 0.5; // Width of the square-well potential
const double max_translation = 0.400; // Å
const double temperature_kelvin = 130.0;
const double sigma = 3.624; // 3.624 Å(borrowed from Kr)
const double specific_epsilon = -0.500; // For specific interaction between Ar and Br (kcal mol-1)
const double nonspecific_epsilon = 0.8; // For nonspecific interactions borrowed from Kr (kcal mol-1)
const double k_boltz = 1.98720650419155E-3; // kcal mol-1 K-1
const int total_moves = 100000;
const float infty = std::numeric_limits¡float¿::infinity();

ofstream fileout0(”8_Trajectory.xyz”);
ofstream fileout1(”8_Conditions_Results.rtf”);
ofstream fileout2(”8_Energy.rtf”);
ofstream fileout3(”8_Distance.rtf”);

// Structures

struct position
{
    double x;
    double y;
    double z;
};

position artificial;
position brtificial;
```

```
position nonspecific[10];
position temporary_position;

// Function Declarations

void record_conditions_results();
void position_initialization();
void random_translation_scheme();
double energy_calculation();
int random_number_generator_atom();
int random_number_generator_select();
double random_number_generator_motion();
double random_number_generator_scheme();

int main()
{
    record_conditions_results();
    position_initialization();
    random_translation_scheme();

    return 0;
}

void record_conditions_results()
{
    fileout1 << ”————(CONDITIONS ONLY)————” << endl;
    fileout1 << ”Cavity Radius (A): ” << cavity_radius << endl;
    fileout1 << ”Maximum Translation (A): ” << max_translation << endl;
    fileout1 << ”Temperature (K): ” << temperature_kelvin << endl;
    fileout1 << ”Sigma (A): ” << sigma << endl;
    fileout1 << ”Specific Epsilon (kcal/mol): ” << specific_epsilon << endl;
    fileout1 << ”Nonspecific Epsilon (kcal/mol): ” << nonspecific_epsilon << endl;
    fileout1 << ”Boltzmann Constant (kcal/mol.K): ” << k_boltz << endl;
    fileout1 << ”Total Moves: ” << total_moves << endl;
    fileout1 << ”—————(COMPLETE!)—————” << endl;
}

void position_initialization()
{
    fileout0 << 2+ number_of_nonspecific_particles << endl;
    fileout0 << ”\n”;
    // Positioning Artificial, Brtificial Particles

    artificial.x = -7.0000;
    artificial.y = 0;
```

```
        artificial.z = 0;

        brtificial.x = 7.0000;
        brtificial.y = 0;
        brtificial.z = 0;

        fileout0 << ”Ar” << ”\t” << artificial.x << ”\t” << artificial.y << ”\t” << artificial.z << endl;
        fileout0 << ”Br” << ”\t” << brtificial.x << ”\t” << brtificial.y << ”\t” << brtificial.z << endl;

        // Positioning Nonspecific Particles

        int m = 0;

        temporary_position.x = -9;
        for (int i = 0; i < lattice_size; i++)
        {
            temporary_position.x + = lattice_distance;
            temporary_position.y = -9;
            for (int j = 0; j < lattice_size; j++)
            {
                temporary_position.y + = lattice_distance;
                temporary_position.z = -9;
                for (int k = 0; k < lattice_size; k++)
                {
                    temporary_position.z + = lattice_distance;

                    nonspecific[m].x = temporary_position.x;
                    nonspecific[m].y = temporary_position.y;
                    nonspecific[m].z = temporary_position.z;

                    m++;
                }
            }
        }
        for (int i = 0; i < number_of_nonspecific_particles; i++)
        {
            fileout0 << ”Kr \t” << nonspecific[i].x << ”\t” << nonspecific[i].y << ”\t” << nonspecific[i].z << endl;
        }
    }

    void random_translation_scheme()
    {
```

```
int n_accept = 0;
int n_reject = 0;

double initial_energy = energy_calculation();
fileout2 << ”Step #0: ” << initial_energy << endl;

for (int move = 1; move <= total_moves; move++)
{
cout << ”Step #” << move << endl;
    // Pick a random (int) number between 0 and 9 for selecting Ar/Br/Kr
    int rndnum = random_number_generator_select();
    // Pick a random (int) number between 0 and 7 for selecting one of Kr’s
    int m = random_number_generator_atom();

    // Calculate old energy
    double old_energy = energy_calculation();

    // Save old position
    double old_information[9] = { artificial.x, artificial.y, artificial.z, brtificial.x, brtificial.y, brtificial.z, nonspecific[m].x, nonspecific[m].y, nonspecific[m].z };

    // Make the move
    double delta_x = random_number_generator_motion();
    double delta_y = random_number_generator_motion();
    double delta_z = random_number_generator_motion();

    // Choose Artificial if rndnum = 8 & Brtificial if rndnum = 8 & Nonspecific if rndnum = 0˜7
    if (rndnum == 9)
    {
        artificial.x = artificial.x + delta_x;
        artificial.y = artificial.y + delta_y;
        artificial.z = artificial.z + delta_z;

        cout << ”Ar ” << ” Chosen: dx = ” << delta_x << ”\t dy = ” << delta_y << ”\t dz = ” << delta_z << endl;
        cout << ”Random Translation Complete!” << endl;
    }
    else if (rndnum == 8)
    {
        brtificial.x = brtificial.x + delta_x;
        brtificial.y = brtificial.y + delta_y;
        brtificial.z = brtificial.z + delta_z;
```

```
            cout << ”Br ” << ” Chosen: dx = ” << delta_x << ”\t dy = ” << delta_y <<
”\t dz = ” << delta_z << endl;
            cout << ”Random Translation Complete!” << endl;
        }
        else
        {
            nonspecific[m].x = nonspecific[m].x + delta_x;
            nonspecific[m].y = nonspecific[m].y + delta_y;
            nonspecific[m].z = nonspecific[m].z + delta_z;

            cout << ”Kr ” << m << ” Chosen: dx = ” << delta_x << ”\t dy = ” << delta_y
<< ”\t dz = ” << delta_z << endl;
            cout << ”Random Translation Complete!” << endl;
        }

        // Calculate new energy

        double new_energy = energy_calculation();

        // Compute the distance from the origin

        double distar = sqrt(artificial.x * artificial.x + artificial.y * artificial.y + artificial.z
        * artificial.z);
        double distbr = sqrt(brtificial.x * brtificial.x + brtificial.y * brtificial.y + brtificial.z
        * brtificial.z);
        double distkr = sqrt(nonspecific[m].x * nonspecific[m].x + nonspecific[m].y
        * nonspecific[m].y + nonspecific[m].z * nonspecific[m].z);

        cout << ”Old Energy: ” << old_energy << ”\t” << ”New Energy: ” << new_energy
<< ”\t” << ”Distances: ” << distar << ” ” << distbr << ” ” << distkr << endl;
        cout << ”\n”;

        bool accept = false;

        // Boundary Condition and Pauli Exclusion
        if ((distar > cavity_radius) || (distbr > cavity_radius) || (distkr > cavity_radius) ||
        (new_energy == infty))
        {
            accept = false;
            n_reject + = 1;
            cout << ”Rejected222!” << ”\n” << endl;
            artificial.x = old_information[0];
            artificial.y = old_information[1];
            artificial.z = old_information[2];
            brtificial.x = old_information[3];
```

```
            brtificial.y = old_information[4];
            brtificial.z = old_information[5];
            nonspecific[m].x = old_information[6];
            nonspecific[m].y = old_information[7];
            nonspecific[m].z = old_information[8];
        }
        else
        {
            // Automatically accept when energy goes down
            if (new_energy <= old_energy)
            {
                accept = true;
                n_accept + = 1;
                cout << ”Accepted111!” << ”\n” << endl;
                fileout0 << 2 + number_of_nonspecific_particles << endl;
                fileout0 << ”\n”;

                fileout0 << ”Ar” << ”\t” << artificial.x << ”\t” << artificial.y << ”\t” <<
artificial.z << endl;
                fileout0 << ”Br” << ”\t” << brtificial.x << ”\t” << brtificial.y << ”\t” <<
brtificial.z << endl;
                for (int i = 0; i < number_of_nonspecific_particles; i++)
                {
                    fileout0 << ”Kr \t” << nonspecific[i].x << ”\t” ¡¡ nonspecific[i].y <<
”\t” << nonspecific[i].z << endl;
                }
                fileout2 << ”Step #” << move << ”: ” << new_energy << endl;

                double eq_dist = sqrt((brtificial.x−artificial.x) * (brtificial.x−artificial.x) +
                (brtificial.y−artificial.y) * (brtificial.y−artificial.y) +
                (brtificial.z−artificial.z) * (brtificial.z−artificial.z));

                if (eq_dist <= sigma + well_width)
                {
                    fileout3 << eq_dist << endl;
                }
            }
            else
            {
                // Test exp(-dE/kT) >= rand[0,1]
                double p = exp((old_energy - new_energy) / (k_boltz*temperature_kelvin));
                double q = random_number_generator_scheme();
                if (p >= q)
                {
                    accept = true;
```

```
            n_accept + = 1;
            cout << ”Accepted222!” << ”\n” << endl;
            fileout0 << 2 + number_of_nonspecific_particles << endl;
            fileout0 << ”\n”;

            fileout0 << ”Ar” << ”\t” << artificial.x << ”\t” << artificial.y << ”\t” << artificial.z << endl;
            fileout0 << ”Br” << ”\t” << brtificial.x << ”\t” << brtificial.y << ”\t” << brtificial.z << endl;
            for (int i = 0; i < number_of_nonspecific_particles; i++)
            {
                fileout0 << ”Kr \t” << nonspecific[i].x << ”\t” << nonspecific[i].y << ”\t” << nonspecific[i].z << endl;
            }
            fileout2 << ”Step #” << move << ”\t” << new_energy << endl;

            double eq_dist = sqrt((brtificial.x−artificial.x) * (brtificial.x−artificial.x)
            + (brtificial.y−artificial.y) * (brtificial.y - artificial.y)
            + (brtificial.z−artificial.z) * (brtificial.z−artificial.z));

            if (eq_dist <= sigma + well_width)
            {
                fileout3 << eq_dist << endl;
            }
        }
        else
        {
            accept = false;
            n_reject + = 1;
            accept = false;
            n_reject + = 1;
            cout << ”Rejected111!” << ”\n” << endl;
            artificial.x = old_information[0];
            artificial.y = old_information[1];
            artificial.z = old_information[2];
            brtificial.x = old_information[3];
            brtificial.y = old_information[4];
            brtificial.z = old_information[5];
            nonspecific[m].x = old_information[6];
            nonspecific[m].y = old_information[7];
            nonspecific[m].z = old_information[8];
        }
    }
}
```

```
            if ((move % 1000) == 0)
            {
                fileout1 << ”Step ” << move << ”\t” << ”Energy (kcal/mol): ” << setprecision(2) << fixed << old_energy
                << ”\t% Acceptances: ” << double(n_accept * 100 / move)
                << ”\t% Rejections: ” << double(n_reject * 100 / move) << endl;
            }
        }
        cout << ”Total Number of Acceptances: ” << n_accept << endl;
        cout << ”Total Number of Rejections: ” << n_reject << endl;
        cout << ”\n”;
    }

    // Square Well Potential

    double energy_calculation()
    {
        double dxab, dyab, dzab;
        double dxak, dyak, dzak;
        double dxbk, dybk, dzbk;
        double dxkk, dykk, dzkk;
        double interatomic_length_ab;
        double interatomic_length_ak_square;
        double interatomic_length_bk_square;
        double interatomic_length_kk_square;
        double sigma2_divide_r2_ak;
        double sigma2_divide_r2_bk;
        double sigma2_divide_r2_kk;
        double interatomic_energy_arkr = 0;
        double interatomic_energy_brkr = 0;
        double interatomic_energy_krkr = 0;
        double total_potential_energy = 0;
        double potential_energy_arbr = 0;
        double potential_energy_arkr = 0;
        double potential_energy_brkr = 0;
        double potential_energy_krkr = 0;

        // Between Ar and Br

        dxab = brtificial.x−artificial.x;
        dyab = brtificial.y−artificial.y;
        dzab = brtificial.z−artificial.z;

        interatomic_length_ab = sqrt(dxab*dxab + dyab*dyab + dzab*dzab);
```

```
if (interatomic_length_ab < sigma)
{
    potential_energy_arbr = infty;
}
else if ((interatomic_length_ab >= sigma) && (interatomic_length_ab <= sigma + well_width))
{
    potential_energy_arbr = specific_epsilon;
}
else
{
    potential_energy_arbr = 0;
}

// Between Ar and Kr

for (int i = 0; i < number_of_nonspecific_particles; i++)
{
    dxak = artificial.x−nonspecific[i].x;
    dyak = artificial.y−nonspecific[i].y;
    dzak = artificial.z−nonspecific[i].z;

    interatomic_length_ak_square = dxak*dxak + dyak*dyak + dzak*dzak;

    sigma2_divide_r2_ak = sigma*sigma / interatomic_length_ak_square;
    interatomic_energy_arkr = 4 * nonspecific_epsilon * (pow(sigma2_divide_r2_ak, 6)
    −pow(sigma2_divide_r2_ak, 3));

    potential_energy_arkr += interatomic_energy_arkr;
}

// Between Br and Kr

for (int j = 0; j < number_of_nonspecific_particles; j++)
{
    dxbk = brtificial.x−nonspecific[j].x;
    dybk = brtificial.y−nonspecific[j].y;
    dzbk = brtificial.z−nonspecific[j].z;

    interatomic_length_bk_square = dxbk*dxbk + dybk*dybk + dzbk*dzbk;

    sigma2_divide_r2_bk = sigma*sigma / interatomic_length_bk_square;
    interatomic_energy_brkr = 4*nonspecific_epsilon*(pow(sigma2_divide_r2_bk, 6)
    −pow(sigma2_divide_r2_bk, 3));

    potential_energy_brkr + = interatomic_energy_brkr;
```

```
        }

        // Between Kr and Kr

        for (int r = 0; r < number_of_nonspecific_particles−1; r++)
        {
            for (int t = r + 1; t < number_of_nonspecific_particles; t++)
            {
                dxkk = nonspecific[r].x−nonspecific[t].x;
                dykk = nonspecific[r].y−nonspecific[t].y;
                dzkk = nonspecific[r].z−nonspecific[t].z;

                interatomic_length_kk_square = dxkk*dxkk + dykk*dykk + dzkk*dzkk;

                sigma2_divide_r2_kk = sigma*sigma / interatomic_length_kk_square;
                interatomic_energy_krkr = 4*nonspecific_epsilon*(pow(sigma2_divide_r2_kk, 6)
                −pow(sigma2_divide_r2_kk, 3));

                potential_energy_krkr + = interatomic_energy_krkr;
            }
        }
        total_potential_energy + = potential_energy_arbr + potential_energy_arkr
        + potential_energy_brkr + potential_energy_krkr;

        return total_potential_energy;
    }

    int random_number_generator_atom()
    {
        uniform_real_distribution⟨double⟩ dist(-0.5, number_of_nonspecific_particles−0.5); // Type
of random number distribution
        mt19937 rng; // Mersenne Twister: Good quality random number generator
        rng.seed(random_device{}()); // Initialize with non-deterministic seeds

        int atom = int(dist(rng));

        return atom;
    }

    int random_number_generator_select()
    {
        uniform_real_distribution⟨double⟩ dist(-0.5, 2 + number_of_nonspecific_particles−0.5);
// Type of random number distribution
        mt19937 rng; // Mersenne Twister: Good quality random number generator
        rng.seed(random_device{}()); // Initialize with non-deterministic seeds
```

```
    int atom = int(dist(rng));

    return atom;
}

double random_number_generator_motion()
{
    double max = max_translation;
    double min = -max_translation;

    uniform_real_distribution⟨double⟩ dist(min, max); // Type of random number distribution
    mt19937 rng; // Mersenne Twister: Good quality random number generator
    rng.seed(random_device{}()); // Initialize with non-deterministic seeds

    double motion = dist(rng);

    return motion;
}

double random_number_generator_scheme()
{
    uniform_real_distribution⟨double⟩ dist(0, 1); // Type of random number distribution
    mt19937 rng; // Mersenne Twister: Good quality random number generator
    rng.seed(random_device{}()); // Initialize with non-deterministic seeds

    double scheme = dist(rng);

   return scheme;
}
```

Appendix E

Copyright Permission

E.1 Permission from the Royal Society of Chemistry

From: Gill Cockhead
Sent: Friday, March 9, 2018 at 2:23 am
To: Myong In Oh
Subject: RE: Permission to Use Copyrighted Material in a Doctoral Thesis

Dear Myong In Oh

The Royal Society of Chemistry (RSC) hereby grants permission for the use of your paper(s) specified below in the printed and microfilm version of your thesis. You may also make available the PDF version of your paper(s) that the RSC sent to the corresponding author(s) of your paper(s) upon publication of the paper(s) in the following ways: in your thesis via any website that your university may have for the deposition of theses, via your universitys Intranet or via your own personal website. We are however unable to grant you permission to include the PDF version of the paper(s) on its own in your institutional repository. The Royal Society of Chemistry is a signatory to the STM Guidelines on Permissions (available on request).

Please note that if the material specified below or any part of it appears with credit or acknowledgement to a third party then you must also secure permission from that third party before reproducing that material.

Please ensure that the thesis states the following:

Reproduced by permission of The Royal Society of Chemistry

and include a link to the paper on the Royal Society of Chemistrys website.

Please ensure that your co-authors are aware that you are including the paper in your thesis.

Regards

Gill Cockhead
Publishing Contracts & Copyright Executive
Royal Society of Chemistry

E.2 Permission from the American Chemical Society

From: EIC Office Analytical Chemistry
Sent: Monday, March 12, 2018 at 9:12 am
To: Myong In Oh
Subject: RE: Permission to Use Copyrighted Material in a Doctoral Thesis

Dear Myong In Oh:

You have my permission to use this article Advances in modeling the stability of non-covalent complexes in charged droplets with applications in ESI-MS experiments in your thesis as long as the correct citations are made as directed in the ACS Thesis/Dissertation Policy and the ACS Journal Publishing Agreement.

Sincerely,

Prof. Jonathan V. Sweedler

Editor-in-Chief

Analytical Chemistry

From: EIC Office Analytical Chemistry
Sent: Friday, March 9, 2018 at 10:54 am
To: Myong In Oh
Subject: RE: Permission to Use Submitted Material in a Doctoral Thesis

Dear Myong In Oh:

You have my permission to use this article “Exploring the extrusion mechanisms of proteins from droplets and of “droplets from macroions in your thesis as long as the correct citations are made as directed in the ACS Thesis/Dissertation Policy and the ACS Journal Publishing Agreement.

Sincerely,

Prof. Jonathan V. Sweedler

Editor-in-Chief

Analytical Chemistry

From: EIC Office Journal of Physical Chemistry
Sent: Friday, March 9, 2018 at 10:32 am
To: Myong In Oh
Subject: RE: Permission to Use Copyrighted Material in a Doctoral Thesis

Greetings;

Permission is granted with the understanding that the appropriate citations of the published work must be made.

Sincerely,

Davine for

George C. Schatz

Editor-in-Chief

E.3 Permission from Springer Nature

From: Journalpermissions
Sent: Friday, March 9, 2018 at 5:53 am
To: Myong In Oh
Subject: RE: Permission to Use Copyrighted Material in a Doctoral Thesis

Dear Myong,

Thank you for your Springer Nature request. As an author, you are welcome to reuse this content in other works created by yourself, but we require a license for our records. Unfortunately it is not our policy to allow our content to be uploaded to any commercial repository and we cannot grant permission for this. Im sorry that we cannot help on this occasion.

To obtain your permission, please access the following URL which will take you to the order entry page from which you can request your permission. Please apply for the thesis/dissertation license, this is a free license. Please note: we allow the use of the final accepted manuscript version of the article in the thesis, not the published version.

```
https://s100.copyright.com/AppDispatchServletpublisherName=Springer
Nature&orderBeanReset=true&orderSource=SpringerLink&author=Myong+In+
Oh\%2C+Styliani+Consta&issueNum=11&contentID=10.1007%2Fs13361-017-175
4-4&openAccess=false&endPage=2279&publicationDate=2017&startPage=226
2&volumeNum=28&title=Charging+and+Release+Mechanisms+of+Flexible+Mac
romolecules+in+Droplets&imprint=American+Society+for+Mass+Spectromet
ry&publication=1044-0305
```

During the process, you will set up an account with RightsLink, who will invoice you if there are any fees incurred. You will be able to use your RightsLink account in the future to request (and pay for) permissions online both from Springer Nature and from other participating publishers. RightsLink will also email you confirmation that permission has been granted, in the form of a printable license.

Best wishes,

Oda Siqveland

Permissions Assistant

SpringerNature

E.4 Permission from the American Institute of Physics

From: AIPRights Permissions
Sent: Wednesday, March 14, 2018 at 11:48 am
To: Myong In Oh
Subject: RE: Permission to Use Copyrighted Material in a Doctoral Thesis

Dear Dr. Oh:

You are permitted to include your published article in your thesis, provided you also include a credit line referencing the original publication.

Our preferred format is (please fill in the citation information):

Reproduced from [FULL CITATION], with the permission of AIP Publishing.

Please let us know if you have any questions.

Sincerely,

Susann Brailey

Manager, Rights & Permissions

AIP Publishing

Curriculum Vitae

Myong In Oh

Post-Secondary Education and Degrees:	Ph.D. in Chemistry (2018) *The* University *of* Western Ontario, London, Ontario, Canada Supervisor: Prof. Styliani Consta
	B.Sc. in Chemistry *with Distinction* (2013) *The* University *of* Western Ontario Supervisor: Prof. Zhifeing Ding
Related Work Experience:	Visiting Graduate Student (2016) *The* University *of* Cambridge, Cambridge, United Kingdom Supervisor: Prof. Daan Frenkel
	Graduate Teaching Assistant (2013-2015) *The* University *of* Western Ontario - Microscopic Phenomena (Winter 2014, Winter 2015) - Discovering Chemical Structure (Fall 2014) - General Chemistry for Engineers (2013-2014) - Chemistry of the Environment (Fall 2013)
	Undergraduate Research Assistant (2012) *The* University *of* Western Ontario Supervisor: Prof. François Lagugné-Labarthet
	Undergraduate Research Assistant (2011) *The* University *of* Western Ontario Supervisor: Prof. Zhifeing Ding
Honours and Awards:	Marie Sklodowska-Curie Actions Seal of Excellence European Commission (2018) International

Alexander Graham Bell Canada Graduate Scholarships - Doctoral Program
Natural Sciences and Engineering Research Council (2015-2018)
National, 105000 CAD

Doctoral Research Excellence Award
The University *of* Western Ontario (2016-2018)
Institutional, 8333 CAD

Canada Graduate Scholarships - Michael Smith Foreign Study Supplements
Natural Sciences and Engineering Research Council (2016)
National, 6000 CAD

CSC Travel Award
The University *of* Western Ontario (2016)
Institutional, 800 CAD

ASPIRE Award
The University *of* Western Ontario (2015)
Institutional, 850 CAD

Alexander Graham Bell Canada Graduate Scholarships - Masters Program
Natural Sciences and Engineering Research Council (2015-2018)
National, 17500 CAD

Doctoral Admission Scholarship
The University *of* Western Ontario (2013)
Institutional, 8000 CAD

Christian Siverts Scholarship
The University *of* Western Ontario (2013)
Institutional, 3500 CAD

Dean's Honor List
The University *of* Western Ontario (2010-2013)
Institutional

Andrew and Sarah Hamilton Scholarship
The University *of* Western Ontario (2010-2013)
Institutional, 4500 CAD

UWO In-Course Scholarship Year IV
The University *of* Western Ontario (2012)
Institutional, 700 CAD

Western Undergraduate Student Research Award
The University *of* Western Ontario (2012)
Institutional, 3200 CAD

NSERC Undergraduate Student Research Award
Natural Sciences and Engineering Research Council (2011)
Institutional, 6100 CAD

Entrance Scholarship of Excellence
The University *of* Western Ontario (2009)
Institutional, 2000 CAD

Publications:

1. **Oh, M. I.**, Consta, S. Potential of mean force of the interactions of two protein-mimicking particles in bulk solution. In preparation.

2. **Oh, M. I.**, Consta, S. Effect of the dielectric properties and molecular structure of solvents on the formation of "star"-shaped droplets. In preparation.

3. **Oh, M. I.**, Consta, S. General solvation motifs of a charged linear macroion in droplets. In preparation.

4. Consta, S., **Oh, M. I.** Relation of solvation and charge state of macroions in droplets: general motifs and universal parameters. *J. Am. Soc. Mass Spectrom.* (Invited critical insight) (submitted on April 1, 2018, Manuscript ID: ASMS-D-18-00107) (25 pages)

5. Malevanets, A., **Oh, M. I.**, Sharawy, M., Consta, S. Landau-Ginzburg theory for accounting for "star"-shaped droplets. *Mol. Phys.* (Invited article to honour Prof. D. Frenkel) (under revision, submitted on March 23, 2018, Manuscript ID: TMPH-2018-0221) (15 pages)

6. **Oh, M. I.**, Consta, S. Exploring the extrusion mechanism of proteins from droplets and of "droplets" from macroions. *Anal. Chem.* (submitted on February 28, 2018, Manuscript ID: ac-2018-00943b) (15 pages + 9 pages of supporting information)

7. Consta, S., **Oh, M. I.**, Sharawy, M., Malevanets, A. Macroion-solvent interactions in charged droplets. *J. Phys. Chem. A* (Invited feature article, Cover) (accepted on March 16, 2018, Manuscript ID: jp-2018-01404v) (12 journal pages)

8. Consta, S., Malevanets, A., **Oh, M. I.**, Sharawy, M. Role of reaction coordinate in free energy calculations. *Mol. Simul.* 2018, 1–11. (in press, DOI:10.1080/08927022.2018.1445245) (11 journal pages)

9. **Oh, M. I.**, Paliy, M., Consta, S. "Star" morphologies of charged nanodrops comprised of conformational isomers. *J. Chem. Phys.* **148**, 024307 (2018). (14 journal pages + 10 pages of supporting information)

10. **Oh, M. I.**, Consta, S. What factors determine the stability of a weak protein-protein interaction in a charged aqueous droplet? *Phys. Chem. Chem. Phys.* **19**, 31965–31981 (2017). (17 journal pages + 24 pages of supporting information)

11. **Oh, M. I.**, Malevanets, A., Paliy, M., Frenkel, D., Consta, S. When droplets become stars: charged dielectric droplets beyond the Rayleigh limit. *Soft Matter* **13**, 8781–8795 (2017). (16 journal pages + 10 pages of supporting information)

12. **Oh, M. I.**, Consta, S. Charging and release mechanisms of flexible macromolecules in droplets. *J. Am. Soc. Mass Spectrom.* (Cover) **28**, 2262–2279 (2017). (18 journal pages + 12 pages of supporting information)

13. Consta, S., Sharawy, M., **Oh, M. I.**, Malevanets, A. Advances in modeling the stability of noncovalent complexes in charged droplets with applications in electrospray ionization-MS experiments. *Anal. Chem.* (Perspective) **89**, 8192–8202 (2017). (11 journal pages + 2 pages of supporting information)

14. **Oh, M. I.**, Consta, S. Stability of a transient protein complex in a charged aqueous droplet with variable pH. *J. Phys. Chem. Lett.* **8**, 80–85 (2017). (6 journal pages + 11 pages of supporting information)

15. Consta, S., **Oh, M. I.**, Malevanets, A. New mechanisms of macroion-induced disintegration of charged droplets. *Chem. Phys. Lett.* (Frontiers, Cover) **663**, 1–12 (2016). (12 journal pages)

16. Ding, Z., Jia, F., Love, D., **Oh, M. I.** 2016. Low-cost photovoltaic-active $CuInS_2$ nanocrystal ink prepared by a mild chemical reaction without post-processing. U.S. Patent 9,346,686 (Tech ID: W-12-016), filed December 24, 2012, and issued May 24, 2016.

17. Consta, S., **Oh, M. I.**, Soltani, S. Advances in the theoretical and molecular simulation studies of the ion chemistry in droplets. *Int. J. Mass Spectrom.* **377**, 557–567 (2015). (11 journal pages)

18. Soltani, S., **Oh, M. I.**, Consta, S. Effect of solvent on the charging mechanisms of poly(ethylene glycol) in droplets. *J. Chem. Phys.* **142**, 114307 (2015). (8 journal pages + 5 pages of supporting information) (Soltani, S. is a graduate student.)

Oral Presentations:

1. **Oh, M. I.**, Consta, S. Computational and analytical modeling of droplet-macroion interactions. Chemistry Graduate Student Symposium, Buffalo, New York, United States (international) May 21-23, 2018 (upcoming)

2. **Oh, M. I.**, Consta, S. Stability of a transient protein complex in a charged aqueous droplet with variable pH. Applied Mathematics, Modeling, and Computational Science (AMMCS) Conference, Waterloo, Ontario, Canada (international) August 24, 2017

3. **Oh, M. I.**, Consta, S. Stability of a transient protein complex in an aqueous droplet with variable pH. Canadian Society for Chemistry Conference, Toronto, Ontario, Canada (international) May 28, 2017

4. **Oh, M. I.**, Consta, S. How does the droplet environment affect the stability of a noncovalent protein complex? Symposium on Chemical Physics, Waterloo, Ontario, Canada (international) November 6, 2016

5. **Oh, M. I.**. Markov chain Monte Carlo simulations. University of Western Ontario (ASPIRE), London, Ontario, Canada (institutional) September 28, 2016

6. **Oh, M. I.**, Consta, S. Fragmentation mechanism of a noncovalent protein vomplex in an aqueous droplet. CECAM, Lausanne, Switzerland (international) July 7, 2016

7. **Oh, M. I.**. Stability of a charged droplet in the presence of a macromolecule (March 10, 2016) and Entropy-driven self-assembling colloids (March 16, 2016). University of Cambridge, Cambridge, United Kingdom (presentations given in the group meetings of the Frenkel group)

8. **Oh, M. I.**, Consta, S. Effect of solvent on solvation and sodiation mechanisms of poly(ethylene glycol) in droplets. Applied Mathematics, Modeling, and Computational Science (AMMCS) Conference, Waterloo, Ontario, Canada (international) June 10, 2015

Poster Presentations:

1. **Oh, M. I.**, Consta, S. Computational and analytical modeling of droplet-macroion interactions. Canadian Symposium on Theoretical and Computational Chemistry, Windsor, Ontario, Canada (national) July 15-19, 2018 (upcoming)

2. **Oh, M. I.**, Consta, S. Stability of a transient protein complex in a charged aqueous droplet with variable pH. Statistical Mechanics Meeting, Berkeley, California, United States (international) January 13, 2018 (poster presentation and three-minute oral presentation)

3. **Oh, M. I.**, Consta, S. Stability of a transient protein complex in an aqueous droplet with variable pH. Gordon Research Conference, Ventura, California, United States (international) February 15, 2017

4. **Oh, M. I.**, Consta, S. Stability of a transient protein complex in an aqueous droplet with variable pH. Gordon Research Seminar, Ventura, California, United States (international) February 12, 2017

5. **Oh, M. I.**, Consta, S. Macroion-induced disintegration of charged droplets beyond the conventional mechanisms. CECAM, Lausanne, Switzerland (international) July 6, 2016

6. **Oh, M. I.**, Consta, S. Single flexible linear macromolecule confined in a nanodroplet. Symposium on Chemical Physics, Waterloo, Ontario, Canada (international) November 7, 2015

7. **Oh, M. I.**, Consta, S. Charging and ejection mechanisms of poly(ethylene glycol) in charged aqueous droplets. Symposium on Chemical Physics, Waterloo, Ontario, Canada (international) November 8, 2014
8. **Oh, M. I.**, Consta, S. Solvation and charging of a macromolecule in neutral and charged aqueous droplets. Canadian Symposium on Theoretical and Computational Chemistry, Montreal, Quebec, Canada (national) July 14, 2014
9. **Oh, M. I.**, Consta, S. Release mechanism of a poly(ethylene glycol) from a highly charged aqueous droplet. Symposium on Chemical Physics, Waterloo, Ontario, Canada (international) November 2, 2013
10. **Oh, M. I.**, Nimley, C., Jia, F., Ding, Z. Synthesis and photoelectrochemistry of $CuInS_2$ nanocrystals for solar cells. CAMBR Distinguished Lecturer and Research Day, London, Ontario, Canada (national) June 23, 2011

Course Work:

- Quantum Mechanics (Final Mark: 91/100)
- Computational Quantum Chemistry (Final Mark: 84/100)
- Molecular Simulations (Final Mark: 93/100)
- Electrostatics in Chemistry (Final Mark: 94/100)
- Seminar (Final Mark: 84/100)
- Advanced Calculus I (undergraduate level, Final Mark: 96/100)
- Advanced Calculus II (undergraduate level, Final Mark: 96/100)